

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(11) International Publication Number:

WO 00/64247

(51) International Patent Classification 7: A01K 67/027, C12N 9/16

A1

(43) International Publication Date:

2 November 2000 (02.11.00)

(21) International Application Number:

PCT/CA00/00430

(22) International Filing Date:

20 April 2000 (20.04.00)

(30) Priority Data:

60/130,508

23 April 1999 (23.04.99)

US

(71) Applicant (for all designated States except US): UNIVERSITY OF GUELPH [CA/CA]; College of Biological Science, Guelph, Ontario N1G 2W1 (CA).

(75) Inventors/Applicants (for US only): FORSBERG, Cecil, W. [CA/CA]; 44 Hands Drive, Guelph, Ontario N1G 2W1 (CA). GOLOVAN, Serguei [CA/CA]; 89-252 Stone Road, Guelph, Ontario NIG 2V7 (CA). PHILLIPS, John, P. [US/CA]; General Delivery, Arkell, Ontario NOB 1C0 (CA).

(74) Agents: CHARI, Santosh, K. et al.; Orange & Chari, 4900-55 King Street West, P.O. Box 190, Toronto Dominion Bank Tower, Toronto-Dominion Centre, Toronto, Ontario M5K 1H6 (CA).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, IP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: TRANSGENIC ANIMALS EXPRESSING SALIVARY PROTEINS

The invention provides a transgenic animal having within its genome a transgene construct for gastrointestinal tract specific expression (57) Abstract of a protein. In a preferred embodiment, the protein is a phytase or a homologue thereof. Such proteins may be heterologous and may be specifically expressed in the salivary gland of the animal by operably linking the nucleic acid sequence encoding the protein with regulatory sequence including a salivary gland protein promoter/enhancer. Also provided are methods of expressing and producing proteins using such nucleic acid constructs. Further, antibodies specific to such proteins and immunological diagnostic kits are also provided.

BEST AVAILABLE C

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL AM AT AU AZ BA BB BE BF BG BJ BR CG CH CI CM CN CU CZ DE DK EE	Albania Armenia Anstria Australia Anstralia Azerbaijan Bosnia and Herzegovina Barbados Belgiam Burkina Faso Bulgaria Beanin Brazil Belarus Canada Central African Republic Congo Swizzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Deumark Estonia	ES FI FR GA GB GE GH GN GR HU IE II IS IT JP KE KG KP KR LC LL LK LR	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Baly Japan Kenya Kyngyzstan Democratic People's Republic of Korea Republic of Korea Razakstan Saint Lacia Licelacastein Sri Laska Liberla	LS LT LU LV MC MD MG MK ML MN MR MW MX NE NL NO NZ PL PT RO SE SG	Lesotho Likusnia Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mamitania Malawi Mexico Niger Nethertands Norwsy New Zealand Poland Portugal Romania Russian Federation Sodan Swedea Singapore	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovenia Slovekia Senegal Swaziland Chad Togo Tajikistan Turkey Trinidad and Tobago Ulcraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe
---	---	--	--	---	---	--	---

WO 00/64247 PCT/CA00/00430

TRANSGENIC ANIMALS EXPRESSING SALIVARY PROTEINS

FIELD OF THE INVENTION

5

10

15

20

25

30

The present invention relates to transgenic animals and, more specifically, to animals genetically modified to express a desired protein.

BACKGROUND OF THE INVENTION

Phosphorus is an essential element for the growth of all organisms. In livestock production, phosphorus deficiency has been described as the most prevalent mineral deficiency throughout the world and feed must often be supplemented with inorganic phosphorus in order to obtain desired growth performance of monogastric animals (e.g. pigs, poultry etc.).

Phytic acid, or phytate, (*myo*-inositol 1,2,3,4,5,6-hexakis dihydrogen phosphate) is a major storage form of phosphorus in cereals and legumes, representing 18% to 88% of the total phosphorus content (Reddy *et al.* 1982). The enzyme phytase (*myo*-inositol hexakisphosphate phosphohydrolase) belongs to the group of phosphoric monoester hydrolases: it catalyzes the hydrolysis of phytate (*myo*-inositol hexakis phosphate) to inorganic monophosphate and lower phosphoric esters of *myo*-inositol or, in some cases, free *myo*-inositol. Phytases are classified either as 3-phytases or 6-phytases based on the first phosphate group attacked by the enzyme. 3-phytase is typical for microorganisms and 6-phytase for plants (Cosgrove, 1980).

Phytase is either absent or present at a very low levels in monogastric animals (Bitar and Reinhold 1972; Iqbal et al. 1994). Consequently, dietary phytate is not digested or absorbed from the small intestine and instead is concentrated in fecal material, thereby contributing to phosphorus pollution in areas of intensive livestock production. Runoff from animal farms leads to contamination of rivers and streams. Such runoff has resulted in rapid drops in the oxygen concentration in rivers and streams due to excessive algal growth in water, which, in turn, has led to an increase in the mortality rate of fish and existing flora and fauna. This is becoming a global problem as pig and poultry production is increased (Miner 1999;Mallin 2000). Furthermore, phytic acid is viewed as an anti-nutritional factor because it interacts with essential dietary minerals and proteins limiting the nutritional values of cereals and legumes in man and animals (Harland and Morris 1995).

10

15

20

25

For the above reasons, various attempts have been made to enable animals to utilize available phytate in feed. Such attempts have included production of low phytate plants (Abelson 1999), addition of phytase to the animal feed (Simons et al. 1990) (Stahl et al. 1999) or transformation of the fodder plants to produce the required phytase (Pen et al. 1993, Verwoerd et al. 1995). A combination of these options, the feeding of phytase to poultry receiving low phytate corn has also been tested (Huff et al. 1998). However, these solutions increase the cost of animal production. Also because phytase is an enzyme, it is susceptible to inactivation by heat and moisture and is generally unstable at the high temperatures used for feed pelleting.

The primary phytase used for supplementing animal feeds is from Asperigillus sp.; however, phytases are produced by a large number of plants and microorganisms (Wodzinski and Ullah 1996) (Dvorakova 1998). A phytase produced by Escherichia coli has been reported to exhibit the highest activity of those reported (Wodzinski and Ullah 1996). This phytase from E. coli was initially cloned as an acid phosphatase gene that was designated APPA (Dassa et al. 1990). Greiner et al. (1991; 1993) purified phytase from E. coli and reported that some of the kinetic properties of the acid phosphatase activity of the native phytase of E. coli were similar to those of the APPA-encoded acid phosphatase. However, the authors did not clone the phytase gene to prove that it was identical to APPA gene. We have subsequently cloned, overexpressed and characterized APPA gene, and shown that the E. coli gene APPA codes for a bifunctional enzyme exhibiting both phytase and acid phosphatase activities (Golovan et al. 2000). Phytases exhibit phosphatase activity, however the relative activities differ widely among enzymes (Wodzinski and Ullah 1996).

Therefore, there is a need for an improved method of allowing access by animals to phytase so as to enable efficient phytate metabolism and, thereby reducing phosphate pollution.

In the field of protein production using recombinant methods, one of the associated problems relates to the lack of required glycosylation. Therefore, a method of producing such glycoproteins is also needed.

30 SUMMARY OF THE INVENTION

In one embodiment, the invention provides a transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, the construct including a trangene encoding a protein, the

15

20

25

30

transgene being operably linked to a first regulatory sequence for salivary gland specific expression of the protein.

In another embodiment, the invention provides a transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, the construct including a trangene encoding phytase or a homologue thereof.

In yet another embodiment, the invention provides a method of expressing a protein, the method comprising the steps of:

- a) introducing a transgene construct into a non-human animal embryo such that a non-human transgenic animal that develops from the embryo has a genome that comprises the transgene construct, wherein the transgene construct comprises:
 - i) a transgene encoding the protein, and
 - ii) at least one regulatory sequence for gastrointestinal tract specific expression of the protein,
 - b) transferring the embryo to a foster female; and,
 - c) developing the embryo into the transgenic animal wherein the transgene is produced in the gastrointestinal tract of the animal.

In a further embodiment, the invention provides a transgenic animal adapted for expressing a protein according to the above method. The invention also provides for the progeny of such animal.

In another embodiment, the invention provides a process for producing a protein comprising the steps of:

- a) obtaining saliva containing the protein from a non-human transgenic animal, the animal containing within its genome a transgene construct, wherein the transgene construct comprises:
 - i) a transgene encoding the protein, and
 - ii) at least one regulatory sequence for salivary gland specific expression of the protein, and

extracting the protein from the saliva.

In a further embodiment, the invention provides a method for expressing a phytase or a homologue thereof in a non-human animal, the method comprising:

- a) constructing a nucleic acid sequence including a transgene construct comprising:
 - i) a transgene encoding the phytase or a homologue thereof, and

10

15

20

25

30

ii) at least one regulatory sequence for gastrointestinal tract specific expression of the protein, and

b) transfecting the animal with the nucleic acid sequence; whereby the animal carries within the genome of its somatic and/or germ cells the transgene construct and wherein the animal expresses the phytase or a homologue thereof in its gastrointestinal tract.

In another embodiment the invention provides a nucleic acid molecule comprising a nucleic acid sequence including a gene encoding a protein, the gene being operably linked to at least one regulatory sequence for gastrointestinal tract specific expression of the protein.

In another embodiment the invention provides an antibody specific to the protein expressed by the above nucleic acid sequence and a test kit for immunologically detecting such protein. The invention also provides for hybridomas secreting such antibodies.

In another embodiment the invention provides cells that are transfected with the above nucleic acid sequence.

In another embodiment, the invention provides a method for producing a protein molecule comprising a glycosylated protein secreted in the saliva that exhibits a novel physiological activity. One example of such an activity is phytase.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the preferred embodiments of the invention will become more apparent in the following detailed description in which reference is made to the appended drawings wherein:

Figure 1 is a schematic diagram representing a method for producing the gene construct of the present invention containing the inducible proline-rich protein (PRP) promoter/enhancer. More specifically, Figure 1 is a schematic diagram illustrating the steps in the construction of the transgenes R15/APPA+intron and R15/APPA used for the generation of transgenic mice.

Figure 2 is a schematic diagram representing a method for producing the gene construct of the present invention containing the SV40 promoter. More specifically, Figure 2 is a schematic diagram illustrating the steps in construction of the plasmid containing the transgene SV40/APPA+intron that was introduced by transfection into mammalian cell lines.

Figure 3 is a schematic diagram representing a method for producing the gene

construct of the present invention containing the constitutive parotid secretory protein (PSP)

promoter/enhancer. More specifically, Figure 3 is a schematic diagram illustrating the steps

RECTIFIED SHEET (RULE 91)

10

15

20

25

30

in construction of the transgenes Lama2/APPA that codes for the native AppA phytase and the Lama2/PSP/APPA that codes for the AppA phytase with the PSP signal peptide sequence.

Figure 4 is a schematic diagram of the Lama2-APPA plasmid containing the APPA transgene.

Figure 5 illustrates the nucleic acid sequence of the Lama2/APPA plasmid containing the E. coli APPA gene (SEQ ID NO: 1).

Figure 6 illustrates the PCR results for transformed mice. More specifically, figure 6 is a picture of an agarose gel illustrating APPA PCR products from genomic tail DNA of third generation offspring from the transgenic female founder mouse 3-1 generated using the Xho1 and Not1 fragment of the Lama2/APPA construct. A second generation phytase gene positive male was crossed with each of two phytase positive transgenic females 9f and 11f (Table 3). From litter 18m x 9f offspring 3, 4, 5 & 6 are PCR positive and from litter 18m x 11f offspring 2 and 3 are PCR positive. Std is the oligonucleotide standard and the numbers on the left are the bp sizes of the standard. Lane C is a negative control reaction mixture that lacks a DNA template and appA is a positive control containing an amplified segment of the phytase gene. The primers used were APPA-UP2 and APPA-KPN.

Figure 7 illustrates the PCR results for transformed founder pigs. More specifically, Figure 7 is a picture of an agarose gel illustrating phytase gene PCR products and β -globin PCR products from genomic tail DNA of five founder piglets from litter 167. Std is a 1 kb ladder. Lane 2 using the phytase primer set is positive for the phytase gene, and all of the samples were positive for the β -globin gene. Lane C is a negative control not containing template DNA. The phytase transgene primer set included APPA-UP2 and APPA-KPN gave an expected fragment size of 750 bp. The primer set for the β -globin gene included PIG-BGF and PIG-BRG gives an expected fragment size of 207 bp.

Figure 8 illustrates the PCR results for transgene rearrangement tests. More specifically, Figure 8 is a picture of an agarose gel showing the PCR products of four separate primer sets used to amplify different segments of the transgene introduced into pig 167-02. The Std contained a kilobase DNA ladder. The primers used included lane 1, APPA-UP2 and APPA-KPN (750 bp); lane 2, APPA -MATURE and APPA-KPN (1235 bp); lane 3 APPA MATURE and APPA-DOWN2 (608 bp); lane 4, PIG-BGF and PIG-BGR (207 bp). lane 5, a negative control without DNA template added; lane 6, the appA gene & primers APPA-UP2 and APPA-KPN. The numbers on the left indicate the sizes of the bands in the standard. No PCR products were detected in the absence of either DNA template or primers.

10

15

20

25

30

Figure 9 illustrates weight and salivary phytase activity of the transgenic boar 167-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 167-02, •; Average weight ± SD of four penmates, •; phytase activity of 167-02, •; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

Figure 10 illustrates weight and salivary phytase activity of the transgenic boar 282-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 282-02, •; Average weight ± SD of five penmates, •; phytase activity of 282-02, •; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

Figure 11 illustrates weight and salivary phytase activity of the transgenic boar 282-04 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 282-04, •; Average weight ± SD of five penmates, •; phytase activity of 282-04, •; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

Figure 12 illustrates weight and salivary phytase activity of the transgenic boar 405-02 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 405-02, •; Average weight ± SD of four penmates, •; phytase activity of 405-02, •; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

Figure 13 illustrates weight and salivary phytase activity of the transgenic boar 421-06 and average weight of the pen-mates at intervals during growth. Symbols: Weight of 421-06, ♠; Average weight ± SD of four penmates, ♠; phytase activity of 421-06, ■; Phytase specific activity, □. Arrows indicate sampling for fecal phosphorus concentration.

Figure 14 illustrates the PCR results of first generation pigs. More specifically, Figure 14 is a picture of an agarose gel showing the PCR analysis of eight liter 154 piglets. The phytase transgenic boar 167-02 was used to breed a non-transgenic female. Std, 100 bp ladder, numbers on left are the sizes of the fragments in each band in bp; lane 167-02, DNA from boar 167-02 1, DNA from 167-02; lane C, is a lane without added DNA; lanes 1-8, are amplified DNA inserts from each of the offspring piglets of the litter. Phytase primers were Lama-UP and APPA-DOWN4. β-globin primers were PIG-BGF and PIG-BGR.

Figure 15 illustrates a sodium dodecylsulfate gel stained with silver demonstrating the sizes of the *E. coli* produced APPA phytase and the APPA phytase produced by the pig and a demonstration that the pig phytase is glycosylated. More specifically, Figure 15 is a picture of a sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) profile of the purified AppA phytase produced in *E. coli* and the purified pig salivary phytase stained directly with silver (A) and a transfer from a similar SDS-PAGE gel transferred to

10

15

20

25

30

nitrocellulose and stained for glyoproteins (B). Creatinase is not glycosylated while transferring is glycosylated. The numbers on the left are the masses in of the molecular mass standards (Std) expressed in kDa.

Figure 15B is a picture of Western blot of the untreated pig AppA phytase and the same phytase after treatment with a combination of three deglycosylating enzymes. Lane 1, Purified AppA phytase produced in E. coli (untreated); lane 2, purified pig phytase (untreated); lane 3, purified pig phytase treated with the combination of deglycosylating enzymes including N-glycosidase F, O-glycosidase and neuraminidase.

Figure 16 illustrates a Western blot of the pig phytase and the *E. coli* produced APPA phytase using monoclonal antibodies directed to the APPA phytase documenting that they have homologous epitopes. More specifically, Figure 6 is a Western blot of the AppA phytase from pig saliva after various purification steps and of purified phytase produced in *E. coli*. A monoclonal antibody prepared against the *E. coli* phytase was used as the primary antibody for detection. Lane 1, saliva from non-transgenic pig 164-04; lane 2, saliva from transgenic pig 167-02; Lane 3, saliva fraction not bound to DEAE-Sepharose; lane 4, salivary phytase bound to DEAE-Sepharose and released with an NaCl gradient; lane 5, salivary phytase further purified by Chromatofocusing with a pH gradient of 4 to 7; lane 6, phytase purified from *E. coli*. The numbers on the left are the masses of molecular mass standards (not shown) expressed in kDa.

Figure 17 illustrates an SDS-Page of the *E. coli* APPA phytase, saliva samples from phytase negative and positive pigs and mice and a corresponding Western blot documenting that phytases from all three sources have homologous antigenic epitopes, but the animal phytases are larger than that produced in *E. coli*. More specifically, Figure 6 is a SDS-PAGE profile of the purified *E. coli* produced AppA phytase and the AppA phytases produced by pigs and mice stained with silver (A) and a Western blot of an identical set of protein samples (B). A polyclonal antibody prepared against the *E. coli* phytase was used as the primary antibody for detection. Lane 1, Purified AppA phytase produced in *E. coli*; lane 2, Saliva from a non-transgenic pig 164-01; lane 3, Saliva from a AppA producing transgenic pig 167-02; lane 4, Purified phytase from pig 167-02; lane 5, Saliva from a non-transgenic mouse; lane 6, Saliva from a transgenic mouse containing R15/APPA transgene induced with isoproterenol; lane 7, Saliva from a transgenic mouse containing the Lama/APPA transgene; Std, Molecular mass markers. The numbers on the left are the masses of molecular mass standards (not shown) expressed in kDa.

20

25

30

Figure 18 illustrates the nucleic acid sequence of the known segment of the R15/APPA + intron plasmid including the vector sequences of pBLCAT3 (SEQ ID NO:2).

Figure 19 illustrates the nucleic acid sequence of the known segment of the R15/APPA + intron transgene construct used for the generation of transgenic mice (SEQ ID NO:3).

Figure 20 illustrates the nucleic acid sequence of the known segment of the R15/APPA plasmid including the vector sequences of pBLCAT3 (SEQ ID NO:4).

Figure 21 illustrates the nucleic acid sequence of the known segment of the R15/APPA transgene construct used for the generation of transgenic mice (SEQ ID NO:5).

Figure 22 illustrates the nucleic acid sequence of the SV40/APPA + intron plasmid (SEQ ID NO:6).

Figure 23 illustrates the nucleic acid sequence of the Lama2/APPA transgene construct used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7).

15 DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the following description, a number of recombinant DNA technology terms are used. The following definitions have been provided in order to enable a clearer understanding of the specification and appended claims:

"Promoter" - a DNA sequence generally described as the 5' region of a gene and located proximal to the start codon. The transcription of an adjacent gene is initiated at the promoter region. If a promoter is an inducible promoter then the rate of transcription increases in response to an inducing agent. A constitutive promoter is one that initiates transcription of an adjacent gene without additional regulation.

"Operably Linked" - a nucleic acid sequence is "operably linked" when placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is "operably linked" to a coding sequence if the promoter causes the transcription of the sequence. Generally, operably linked means that the linked nucleic acid sequences are contiguous and, where it is necessary to join two protein coding regions, contiguous and in one reading frame.

"Phytase" - any protein that liberates phosphate from myo-inositolhexakis-phosphate or other inositol phosphates. Its catalytic capability may be limited to phytic acid or one of its salts, or it may show less specificity and hydrolyze a variety of phosphorylated compounds.

10

15

20

25

30

"Gene" - a DNA sequence that contains a template for an RNA polymerase and contains information needed for expressing a polypeptide or protein.

"Polynucleotide Molecule" - a polydeoxyribonucleic (DNA) acid molecule or a polyribonucleic acid (RNA) molecule.

"Expression" - the process by which a polypeptide is produced from a structural gene.

"Cloning vehicle" - is a plasmid or phage DNA or other DNA sequence which is capable of carrying genetic information into a host cell. A cloning vehicle is often characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the vehicle. A cloning vehicle is a DNA sequence into which a desired DNA may be spliced in order to bring about its cloning into the host cell.

"Vector" - is a term also used to refer to a cloning vehicle.

"Plasmid" - is a cloning vehicle generally comprising a circular DNA molecule that is maintained and replicates autonomously in at least one host cell.

"Expression vehicle" - a vehicle or vector similar to a cloning vehicle but which supports expression of a gene that has been cloned into it, after transformation of a host. The cloned gene is usually placed under the control of (i.e. is operably linked to) certain control sequences such as promoter sequences.

"Host" - a cell that is utilized as the recipient and carrier of recombinant material.

"Homologous" - refers to a nucleic acid molecule that originates from the same genus or species as the host.

"Heterologous" - refers to a nucleic acid molecule that originates from a different genus or species than that of the host.

"Glycoprotein" - refers to a peptide molecule that has undergone glycosylation.

"Glycosylation" - refers to the addition of carbohydrate groups to a amino acid

residues of a peptide molecule.

In recent years, transgenic animals have been developed for many purposes (Pinkert et al. 1990) (Wall et al. 1997). One premise, therefore, for the present invention is that by providing a transgenic animal capable of expressing phytase, the problems discussed above would be obviated. The options for heterologous phytase expression in animals include (i) salivary gland secretion of a phytase, (ii) pancreatic secretion of the enzyme into the small intestine along with the digestive enzymes, or (iii) secretion from the intestinal epithelial cells much like that of indigenous alkaline phosphatase and glycosidases (Low, 1989). The E. coli phytase would appear to be best suited for hydrolytic activity in the monogastric stomach

WO 00/64247 PCT/CA00/00430 10

because the enzyme has a pH optimum in the range of 2.5 to 4.5 and it is resistant to pepsin, the predominant protease active in the stomach. The phytase has a periplasmic location in E. coli and has an N-terminal signal peptide sequence (Golovan et al., 1999) that seemed optimally adapted for secretion from the parotid gland. Phytase could be expressed in either the pancreas for secretion into the small intestine or it could be expressed in the intestinal epithelial tissue and secreted into the intestinal milieu. However, if these choices of expression locations were chosen, it would be necessary to select an enzyme active at the more neutral pH of the small intestine and one which was more resistant to pancreatic enzymes including trypsin, chymotrypsin and elastase.

5

10

15

20

25

30

Factors of importance in terms of the expressed enzyme when selecting a phytase for expression in the gastrointestinal tract include a pH that is optimum for activity, high catalytic activity, broad substrate specificity, and protease resistance. If any of these properties, or indeed others, is not acceptable, there are now sophisticated molecular methods for modifying the properties of an enzyme. These include site directed mutagenesis, random mutagenesis and various modifications of DNA shuffling (Harayama, 1998; Crameri et al., 1998).

Synthesis of phytase in the salivary gland and secretion in the saliva would, therefore, provide for early contact of the enzyme with phytic acid present in the feed and provide sufficient time for hydrolysis.

The salivary gland system of the pig consists of three pairs of glands, the parotid gland, which secretes through a duct on each cheek, and mandibular and submaxillary glands that have joint ducts that secrete beneath the front on the tongue. Saliva secreted in the pig via these ducts is discontinuous and is produced during consumption of solid foods, and can equal the weight of food consumed when water is limited during feed consumption (Corring, 1980; Arkhipovets, 1956). For example, the quantity of saliva produced by a 45 kg pig can vary from near zero when the pig receives a mainly liquid diet to 500 g when a dry diet is consumed without access to water. The salivary glands of the pig secrete amylase (Rozhkov and Galimov, 1990) and a variety of other salivary proteins and mucopolysaccharides.

To our knowledge no porcine genes coding for salivary proteins have been cloned. However, genes coding for major proteins secreted by the rat and mouse have been cloned and characterized. A multigene family encoding a group of unique proteins high in proline, the socalled proline-rich proteins (PRPs) are produced when either mice or rats consume tannins or are injected with isoproterenol.

10

15

20

25

It would be advantageous to develop an animal that is transformed to express phytase, preferably in the salivary gland. In such case, the phytate naturally occurring in the animal feed can be utilized by the animal without any additives being used. This will decrease the cost of animal production, and furthermore, will avoid polluting the environment with phosphorus. Therefore, the present invention aims to overcome the deficiencies of the prior art relating to increasing phytate utilization and, particularly, to provide transgenic animals which express phytase.

In the production of heterologous proteins by means of recombinant methods, several hurdles have been faced. One such hurdle that is often faced is the lack of required post-translational modification of the expressed protein thereby resulting in a protein that is structurally and/or functionally, different from the desired molecule. Glycosylation is one such post-translational modification that is desired. However, such modification is generally found to occur in more complex mammalian systems. Therefore in one embodiment of the present invention there is provided a method of producing recombinant glycoproteins.

In one embodiment, the present invention provides an animal capable of inducible or constitutive salivary expression of a heterologous protein. To illustrate this, the mouse was chosen as the animal model and the gene constructs used for transformation were created using the rat proline-rich protein (PRP) promoter/enhancer (inducible promoter) and the mouse parotid secretory protein (PSP) promoter/enhancer (constitutive promoter). In this illustration, phytase was used for expression in saliva.

After finding that an inducible phytase could be expressed in the parotid gland of mice the expression of the phytase transgene under the control of the constitutive PSP promoter was then tested. Two mice transgenic for the PSP construct were produced under contract at the University of Alabama.

Following the testing of the mice described above, transgenic pigs were developed by introduction into the genome a phytase transgene consisting of a constitutive promoter driving the synthesis of a highly active phytase. The pigs so generated were found to excrete less phosphorus in their feces than non-transgenic pigs.

30 Expression in the Salivary Glands

Saliva is a clear colorless fluid secreted by major salivary glands (parotid, submandibular, sublingual and minor salivary) that lubricates and cleans the oral structure, as well as initiates the process of digestion. The parotid glands are two of six major glands associated with the production of saliva. The parotid gland is composed mainly of two cell

RECTIFIED SHEET (RULE 91)

10

15

25

30

types: acinar and interglobular duct cells. The acinar cells, which represent 75 to 85% of the tissue, are the sites of secretory protein synthesis (Frandson and Spurgeon 1992). Two very abundant proteins are produced by these cells: α-amylase (AMY-1) (2% of polyA RNA) (Madsen and Hjorth 1985), and parotid secretory protein (PSP) (10% of polyA RNA) (Shaw and Schibler 1986). Several constructs are now available which allow tissue-specific expression of a transgene in the salivary glands of mice.

The salivary secretion in pigs has not received the attention given to that of mice and humans. It was suggested that salivary secretion is discontinuous (less secreted between periods of meal consumption). Up to 500 g of saliva may be secreted by a 45 kg pig upon consumption of 500 g of dry feed (Corring 1980). Wide variations were detected in both the flow rate and electrolytes in saliva between animals and even between samples taken from the same animal on separate days (Tryon and Bibby 1966). Very little is known about the composition of pig's saliva or salivary enzymes. Salivary amylase was detected, although the quantity was 250 000 times less than that of pancreatic amylase, and 100 times less than in human saliva (Low 1989). There are no constructs known which would allow salivary gland-specific expression of transgene in pigs.

I) APPA Gene Under Control Of An Inducible Promoter

20 1) Construction of R15/APPA constructs (Inducible Promoter)

In this process, a plasmid is constructed by linking a promoter/enhancer for a saliva protein with the APPA gene, which codes for the bifunctional phytase, acid phosphatase. The APPA gene used in this construction was cloned from E. coli ATCC 33965 into pBR322. This is described above (Golovan et al., 2000).

Proteins, unusually high in proline, the so-called proline-rich proteins (PRPs), comprise about 70% of the total proteins in human saliva (Bennick 1982). Unlike the constitutive expression of the PRPs in humans, the salivary glands of mice, rats and hamster normally either do not express PRPs or express them in low levels. In the rat and mouse, PRP gene expression can be dramatically induced by diets high in tannins or by injection with the β-agonist isoproterenol (Carlson 1993). After 6 to 10 days of daily isoproterenol injection the PRPs comprised about 70% of the total soluble protein in parotid gland extracts. PRP cDNA and PRP genes have been cloned and characterized from rats (Clements et al.

10

1985), mice (Ann and Carlson 1985), hamsters (Mehansho et al. 1987), and humans (Kim and Maeda 1986).

Transgenic mice were used to locate the cis-acting DNA elements that are essential for salivary-specific and inducible expression of the rat proline-rich protein gene, R15. It was found that a parotid control region (-6 to -1.7 kb) upstream of the R15 promoter is capable of directing parotid-specific and isoproterenol-inducible expression of a heterologous promoter construct (Tu et al. 1993). The distal -10 to -6 kb region was shown to function as an enhancer, which can increase levels of expression more than 30-fold. The -6 to -1.7 kb region also seems to function as a locus control region (LCR), because it conferred copy number-dependent and chromosomal position-independent expression of a reporter gene in 15 out of 15 independent transgenic mice (Tu, Lazowski, Ehlenfeldt, Wu, Lin, Kousvelari, and Ann 1993).

We obtained the R15-PRP promoter from Dr. D.K. Ann as a plasmid -10R15/CAT, which placed the chloramphenicol acetyltransferase gene (CAT) under control of the inducible R15-PRP promoter. We decided to use the plasmid as a basis for transgene 15 construction (Figure 1). Due to the absence of complete sequence information about the R15-PRP promoter (only 2 kbp out of 10 kbp was sequenced) we removed the R15-PRP promoter by Xho I digestion (Figure 1, step 1). Re-ligated plasmid was used as a template for PCR with CAT-ATG and CAT-TAA synthetic primers. The 4.3 kbp CAT_{PCR} fragment had the initiation site of the CAT gene substituted with the optimal eukaryotic initiation 20 sequence (Kozak 1987). The fragment was purified by agarose gel electrophoresis, re-ligated to itself and used to transform E. coli (Figure 1, step 2). The CATPCR plasmid was digested with Nco I and filled-in using T4 DNA polymerase to generate a blunt end. After that, the CAT_{PCR} fragment was digested with Eco47III and purified by agarose gel electrophoresis (Figure 1, step 3). Three rare codons in the APPA gene were modified during the sub-cloning 25 steps leading to the construction of the transgene. Specifically, the Ala3 coding sequence was changed from GCG to GCC, the Pro₄₂₈ sequence was changed from CCG to CCC, and the Ala429 sequence was changed from GCG to GCT. This modification was made in order to increase the possibility of transcription of the gene in eukaryotic cells. The APPA gene was amplified by PCR using the previously cloned APPA gene from the pBR322/APPA plasmid 30 with the synthetic primers APPA-DRA and APPA-SMA. The 1.3 kbp APPA_{PCR} fragment generated by PCR was digested with Dra I and Sma I and gel-purified (Figure 1, step 4). APPA_{PCR} and CAT_{PCR} fragments were blunt end ligated to produce CAT/APPA+intron

15

20

25

30

vector (Figure 1, step 5), which was introduced into a DH5α strain of *E. coli*. The insert orientation was checked by restriction digest with Sal I and EcoR I. The transgene in CAT/APPA+intron was checked by sequencing both strands. To remove the SV40 small t intron the 2.3 kbp *APPA*/intron/polyA fragment was excised from a plasmid by Xho I and EcoR I digestion (Figure 1, step 6a), gel purified and digested by Dra I (Figure 1, step 6b). The 1.5 kbp (*APPA*) and 0.2 kbp (polyA) fragments were gel-purified and linked together in three way ligation with CAT_{PCR} digested with Xho I and EcoR I (Figure 1, step 6c). The resulting plasmids CAT/APPA and CAT/APPA+intron were digested with Xho I, gel-purified and re-ligated with R15-PRP promoter digested with Xho I (Figure 1, step 7).

Because of the low efficiency of ligation the whole ligation mixture was used to transform *E.coli*, total plasmid DNA was prepared and run on the agarose gel. Plasmids which were larger than the original CAT/APPA (5.6 kbp) were eluted and re-transformed in *E.coli*. Plasmids with the R15-PRP insert (15 kbp) were identified by electrophoresing DNA from a single colony on an agarose gel. The correct orientation was identified by PCR with R15-

UP1 and APPA-DOWN2 synthetic primers. The plasmids R15/APPA and R15/APPA+intron were both digested with Hind III and Kpn I; transgenes were gel-purified and further purified using a Qiagen column (Figure 1, step 8).

Figure 18 illustrates the nucleic acid sequence for the plasmid containing the known segment of the R15/APPA + intron sequence including the vector sequences of pBLCAT3. The sequence of this plasmid is designated as SEQ ID NO:2.

Figure 19 illustrates the nucleic acid sequence for the transgene construct containing the known segment of the R15/APPA + intron sequence used for the generation of transgenic mice. The sequence of this transgene is designated as SEQ ID NO:3.

Figure 20 illustrates the nucleic acid sequence for the plasmid containing the known segment of the R15/APPA sequence including the vector sequences of pBLCAT3. The sequence for this plasmid is designated as SEQ ID NO:4.

The pBLCAT3 sequence indicated above is present in the CAT/APPA of Figure 1 and in the CAT/APPA+intron of Figure 2. This sequence was part of the original -10R15/CAT and a portion of it was carried through in the construction process.

Figure 21 illustrates the nucleic acid sequence for the transgene construct containing the known segment of the R15/APPA sequence used for the generation of transgenic mice.

The sequence of this transgene is designated as SEQ ID NO:5.

15

20

25

30

2) Expression of SV40/APPA+intron in Cell Culture

To produce an SV40/APPA plasmid for expression of APPA in cell culture, the SV40 promoter/enhancer was amplified by PCR from the pSV-β-galactosidase plasmid (Promega) using the synthetic primers SV-HIND and SV-XHO. The SV40 promoter/enhancer fragment was digested with Xho I and Hind III, gel purified, and ligated into CAT/APPA digested with Xho I and Hind III (Figure 2).

Figure 22 illustrates nucleic acid sequence for the SV40/APPA + intron. The sequence for this plasmid is designated as SEQ ID NO:6.

We obtained a rat parotid acinar cell line PARC 5.8 (Quissell et al. 1998) that we intended to use for transient expression of the phytase transgene. The purpose was to test the efficiency of different constructs for transgene expression and also to detect any deleterious effects of phytase expression before introduction into the animals. We tried transient expression of the APPA gene using R15/APPA and R15/APPA+intron constructs but because of low transfection efficiency and/or low expression levels, we were unable to detect either phytase or β-galactosidase that we used as a control for transfection.

We exchanged the R15-PRP inducible promoter from the R15/APPA construct with the SV40 constitutive promoter-enhancer, which enables high level transient expression in different cell cultures. CHO, COS7 and HELA cell lines were screened for transient expression of the APPA phytase using the SV40 promoter/enhancer. All cell lines were maintained on DMEM/F12 (Sigma) cell medium with 10 % (wt/vol) heat-inactivated fetal bovine serum at 37°C in 5% CO2 and 95% air. Cells were grown to 70 % confluence before transfection. Two hours before transfection the medium was exchanged with fresh medium. Cells were transformed with 5 µg of DNA per 60 mm culture plate (1:1 SV40/APPA and SV40/β-galactosidase) using the DNA-Calcium-Phosphate method of transfection (Gorman et al. 1983). After 6 hours of incubation the medium was removed and cells were subjected to glycerol shock for 3 min (Ausbel et al. 1992). Cells were washed with phosphate-buffered saline (PBS) and incubated in fresh medium under standard growth conditions. After 48 hours of incubation cell-free culture fluid was collected, the cells washed two times with PBS and lysed with 1ml of 1% (vol/vol) NP-40, 1mM disodium EDTA in Hanks balanced salts (HBSS) for 1 hour at 40C. The phytase assay was performed in a final volume of 100 μl of 0.1 M sodium acetate/acetic acid buffer (pH 4.5) using sodium phytate (4 mM) as a substrate at 37°C. After 6 hours of incubation the reaction was stopped with 67 µl ammonium molybdate/ammonium vanadate/nitric acid mixture and the concentration of liberated

PCT/CA00/00430 WO 00/64247 16

inorganic phosphate determined at 405 nm (Engelen et al. 1994). One unit (U) of enzyme activity was the amount of the enzyme releasing 1 µmol inorganic phosphate per minute. The assay was performed in triplicate. As a control for endogenous phytase activity, nontransfected cell lines were used.

We did not detect endogenous phytase activity in non-transfected cell lines. Phytase activity was detected in all transfected cell lines, with COS7 cells expressing a total of 0.35 U of phytase in cell-free culture fluid (4 ml) and 0.0034 U in the cell fraction (1.1 ml) obtained from the same plate. The phytase activity produced by COS7 cells was 7 times higher than that of CHO and 35 times more than the HELA cell line. More than 99% of activity was located in cell-free culture fluid, which suggests that the expressed enzyme was exported out of the cell using the bacterial signal sequence. We were unable to detect expression of cytoplasmic β -galactosidase, which we wanted to use as a control for transfection efficiency.

3) Expression of R15-PRP/APPA in Transgenic Mice

5

10

15

20

25

30

Transgenic mice were generated using the constructs R15/APPA and R15/APPA+intron by Dr. C.A. Pinkert at the NICHD Transgenic Mouse Development Facility (NTMDF). University of Alabama at Birmingham, Alabama. The procedures followed in generating the mice have been standardized by the NTMDF and further information concerning this can be obtained at: http://transgenics.bhs.uab.edu/page1.htm, the content of which is incorporated herein by reference. This procedure involved the microinjection technique for transfecting mice with the desired nucleic acid sequence. To summarize, the sequences are microinjected into mouse zygotes and the surviving eggs are implanted into pseudopregnant recipient mice. The recipient mice then give birth to the resulting founder transgenic mice. It will be appreciated that various other methods of generating transgenic mice may be used in the present invention.

The R15/APPA transgene in mice was detected by PCR using the primers CAT-UP1 and APPA-DOWN2 that gives rise to a 700 bp fragment using the standard PCR conditions, except that the hybridization step was set at 51°C for 40 seconds and the polymerization step was at 72°C for one minute.

For the R15/APPA construct 8 PCR positive founder mice were obtained of which 4 were males and 4 were females. Three of the founders did not pass the transgene to progeny and were probably mosaics. For R15/APPA+intron 5 PCR positive founder mice were obtained, 3 were males and 2 were females, and one of them was found to be mosaic. At 10

10

15

. 20

25

30

to 12 weeks of age PRP production in the PCR positive progeny from different lines was induced for 10 days by daily intraperitoneal (ip) injection of 1mg isoproterenol dissolved in 100 µl sterile saline. To serve as a control several PCR negative progeny were also induced. No significant differences in weight were noticed between PCR positive and PCR negative progeny at either the beginning or end of the induction period. Saliva was collected before induction and at the end of the 10 day induction period.

To collect saliva, mice were lightly anesthetized with a ketamine/xylazine mixture (ip injection of 50 mg ketamine and 5 mg xylazine per kg body weight diluted in water) and saliva flow was induced by injection with pilocarpine/isoproterenol (ip injection of 0.5 mg pilocarpine and 2 mg isoproterenol per kg body weight dissolved in saline) (Hu et al. 1992). Between 100-250 µl of saliva was collected from each mouse over a 30 min period beginning 5 min after the pilocarpine/isoproterenol injection.

The saliva was collected from each mouse by holding it in one hand and withdrawing saliva from the corner of the mouth with a 20 μ l pipetter. Collected saliva was transferred to a cold Eppendorf microcentrifuge tube containing 2 μ l of 0.5 M EDTA (pH 8.0) and 4 μ l of 10 mg/ml protease inhibitor Pefabloc (Boehringer Mannheim) dissolved in water. The tubes with saliva were kept on ice until assays were conducted. Phytase activity in the saliva was assayed as described for the SV40/APPA expressed in cell culture.

Phytase expression was not detected in either un-induced or in induced PCR negative mice. For PCR positive mice, phytase expression was not detected in those that were un-induced. However, phytase expression was observed for PCR positive mice that were induced. The results of this study are summarized in Table 1.

Even though it was possible to distinguish saliva from induced PCR positive from that of PCR negative mice in a phytase assay by a characteristic yellow color, saliva from some of the negative mice, when assayed, produced cloudiness that was impossible to remove by centrifugation and that affected spectrophotometer readings. We did not notice any gender differences in expression, both males and females were found to produce phytase in saliva. In three lines (all R15/APPA+intron) no phytase expression or very low level of expression (0.03-0.95 U/ml) was detected, in 4 lines the level of expression ranged from 7 to 87 U/ml, and two lines (both R15/APPA) produced very high levels of phytase in saliva, 252 and 547 U/ml.

These experiments demonstrated that phytase can be expressed at a very high level in the salivary glands of mice, without detrimental effects on the animals. We also were able to

WO 00/64247 PCT/CA00/00430

produce progeny with an inducible salivary phytase from animals expressing the inducible phytase thereby documenting inheritance of the trait, and showing that the reproductive capability of animals was not affected. When the F2 generation of mice were tested for salivary phytase the level of phytase production was preserved.

Founders containing the transgene without the intron gave offspring that produced significantly higher levels of phytase. The SV40 intron in the R15/APPA+intron construct seems to cause a lower level of expression, and in three lines (A1f, A20f and B0m) the level of phytase was barely detectable. The level of phytase expression in A2m line (R15/APPA+intron) was 6.2 times lower than that of the B0m-intron line (R15/APPA).

Preliminary experiments showed that when the enzyme was analyzed by PAGE its size was increased from 42 kDa to 60 kDa. It is likely modified by glycosylation, but stable and active.

II) APPA Gene Under Control Of A Constitutive Promoter

15

20

25

30

10

5

1) Construction of the Lama2/APPA Transgene (Constitutive Promoter)

The murine parotid secretory protein (PSP) is the most abundantly expressed protein in the parotid gland of mice (Madsen and Hjorth 1985). After an hour of pulse labeling, the mouse parotid gland incorporates 65 to 85% of ¹⁴C-leucine into this single protein (Owerbach and Hjorth 1980). It was estimated that PSP mRNA accumulates up to 50,000 molecules per cell and that from 3 to 5 molecules of PSP are produced for every molecule of amylase (Madsen and Hjorth 1985). Despite the predominance of the PSP in saliva its function is not well characterized.

The single-copy gene coding for PSP has been cloned and characterized. It has two alleles PSP³ (Shaw and Schibler 1986) and PSP³ (Owerbach and Hjorth 1980). The PSP³ allele is also expressed in the sublingual gland, but at 1/10 of the level found in the parotid gland. It was shown that 4.6 kbp of 5' flanking sequence of PSP³ is sufficient for salivary gland specific expression. The level of sublingual expression approached 100% of the PSP mRNA level, whereas the parotid expression did not exceed 1% (Mikkelsen et al. 1992), which demonstrates that regulatory sequences for sublingual and parotid expression are not identical. The level of expression was also dependent on the site of integration. The same construct was used for expression of the C-terminal chain of the human blood coagulation factor VIII, FVIII. A high level of FVIII mRNA was detected in the sublingual gland and a low level in the parotid gland. The transgenic lines also secreted the FVIII light chain into

10

15

20

30

saliva at the level of about 10 units per salivation (about 0.05 ml of saliva) (Mikkelsen et al.,1992). Later the same group achieved a high level of parotid-specific expression that was similar or even exceeded that of the endogenous gene by using 11.4 kbp of 5' flanking sequences and 2.5 kbp of 3'flanking sequences (Larsen et al. 1994). The expression also seems to be position-independent and copy-number-dependent that could indicate the presence of a LCR in these sequences.

Lama 2 is a portion of the PSP gene and comprises an 18 kbp construct that is expressed in transgenic mice at up to 56% of the endogenous PSP gene.

Because a large part of Lama 2 had not been sequenced, the construct was first disassembled and subcloned into pBluescript KS(+) and after incorporation of the APPA gene, the Lama 2 was reassembled back (Figure 3). We used unique enzymes RsrII and Smal to remove a 3.4 kbp fragment from Lama2, which was subcloned into the multiple cloning site (MCS) of pBluescript II KS(+) that was previously digested with Kpn1 and Smal, using a Kpn1-RsrII adapter (Figure 3, step 1).

KpnI* RsrII

TGGGAGGTCG

CATGACCCTCCAGCCAG

That allowed us to preserve the RsrII (CG/GWCCG) site and destroy the Kpn1 site (GGTAC/C> GGTAC/<u>T</u>), which would otherwise interfere with future cloning. The pKS/Lama construct was digested with Apa1 and Kpn1 and used in a three-way ligation with the modified APPA (Figure 3, step 2). We designed two PSP/APPA constructs. One construct APPA-signal/APPA (Figure 3, steps 3a-7a) had the original bacterial signal sequence from the APPA protein having the following amino acid sequence:

25 Met-Lys-Ala-Ile-Leu-Ile-Pro-Phe-Leu-Ser-Leu-Leu-Ile-Pro-Leu-Thr-Pro-Gln-Ser-Ala-Phe-Ala

We also modified a sequence near the ATG codon to resemble the optimal mammalian Kozak sequence (GCC GCC A/GCC ATG G) (Kozak 1987), but we did not mutagenize the +4 position because it would change Lys to Glu in the signal sequence with possible deleterious consequences for protein export. This optimized sequence was used in our previous construct R15/APPA and led to high levels of phytase production. We checked the APPA bacterial signal sequence using the PSORT computer neural network trained on eukaryotic signal sequences and further described at http://psort.nibb.ac.jp:8800/ (Nakai and

RECTIFIED SHEET (RULE 91)

Kanehisa 1992). The APPA bacterial signal sequence was recognized as an efficient leader peptide and the cleavage site was correctly predicted. PSORT also predicted that there is a high probability that phytase would be exported correctly outside of the cell. There were also publications showing that some bacterial signal sequences might function efficiently in mammalian cells (Williamson et al. 1994) (Hall et al. 1990). Our experiments using cell culture demonstrated that the APPA signal was correctly processed with export of phytase outside of the cell.

Experiments using cell culture cannot predict the direction of export and if phytase were exported into blood vessels instead of salivary ducts that could lead to deleterious effects. That is why we also designed a second construct PSP-signal/APPA (Figure 3, steps 3b-7b) that would preserve the original PSP signal amino acid sequence:

Met-Phe-Gln-Leu-Gly-Ser-Leu-Val-Val-Leu-Cys-Gly-Leu-Leu-Ile-Gly-Asn-Ser-Glu-Ser

15 This leader peptide was also efficiently recognized by PSORT with the correct cleavage site (Nakai and Kanehisa 1992). In this construct we also preserved the original PSP sequences near the ATG start codons, which may not be optimal, but could be important in regulation of gene expression. The APPA gene for both constructs was amplified by PCR using as the template our previous transgenic construct R15/APPA that possessed the optimal 20 Kozak sequence and the modified codons for residues Ala3, Pro428 and Ala429 as described earlier. For the APPA signal/APPA construct two synthetic primers were used which introduced a Cla1 site near the ATG codon (APPA-CLA) and a Kpn1 site near the TAA stop codon (APPA-KPN). The APPA_{PCR}1 product was digested with Cla1 and Kpn1. The Cla1 site was also introduced into Lama 2 using pKS/Lama 2 as template for PCR. LAMA-UP 25 primer was located upstream of Apa1 site and the LAMA-CLA primer introduced the Cla1 site near ATG codon (Figure 3, step 3a). Lamapor 1 product was digested with Cla1 and Apal (Figure 3, step 4a). pKS/Lama (Apal-Kpnl), Lama_{PCR}1 (Apal-Clal) and APPA_{PCR}1 (Clal-Kpn1) were combined together in a three-way ligation reaction (Figure 3, step 5a). The recovered pKS/Lama/APPA plasmid was digested with RsrII, Smal and inserted back 30 into Lama2 (Figure 3, step 6a).

For the PSPsignal/APPA construct, the synthetic APPA -KPN primer was used with the synthetic APPA -MATURE primer, which produced phytase without a signal sequence. The APPA_{PCR}2 product was blunt-ended using T4 DNA polymerase and digested with Kpn1. The PSP signal sequence was produced using the LAMA-UP and LAMA -SIGNAL primer

10

30

(Figure 3, step 3b). The Lama_{PCR}2 was blunt-ended using T4 DNA polymerase and digested with Apa1 (Figure 3, step 4b). pKS/Lama (Apa1-Kpn1), Lama_{PCR}2 (Apa1-blunt) and APPA_{PCR}2 (blunt-Kpn1) were combined together in a three-way ligation reaction (Figure 3, step 5b). The recovered pKS/Lama/APPA plasmid was digested with RsrII, Sma1 and inserted back into Lama2 (Figure 3, step 6b).

Even though both constructs were successfully produced we decided to use Lama2/APPA signal/APPA for the generation of transgenic mice, because we have results from our previous transgenic constructs R15/APPA and R15/APPA+intron which demonstrated that phytase with optimized Kozak sequence and the APPA signal peptide was synthesized at a high level in salivary glands after induction and was efficiently exported into the salivary duct. The Lama2/APPA vector was digested with XhoI and NotI, and the transgene was gel-purified and further purified using a Qiagen column (Figure 3, step 7a).

2) Sequence of the Lama2/APPA Construct

A large segment of the Lama2 construct (Laursen and Hjorth 1997) used for 15 construction of the Lama2-APPA transgene had not been reported in GenBank prior to our research. To ensure that we could more clearly describe the transgene construct, and furthermore to avoid the introduction of deleterious DNA sequences from the mouse into the pig in the process of generating transgenic pigs, we sequenced the Lama2-APPA plasmid on both strands. Figure 4 illustrates schematically the structure of the Lama2-APPA plasmid. 20 Figure 5 illustrates the nucleic acid sequence (SEQ ID NO:1) of such plasmid. The full transgene sequence was reconstructed from overlapping DNA sequences using the Contig Assembly Program (CAP) (http://hercules.tigem.it/ASSEMBLY/assemble.html) developed by Huang (1996; 1999) and then inspected manually for sequencing errors. The transgene sequence was checked for the presence of interspersed repetitive elements using the computer 25 program RepeatMasker (Smith and Green, RepeatMasker at http://ftp.genome.washington.edu/cgi-bin/RepeatMasker). It was found that 26 % of the transgene sequence was composed of repetitive elements (Table 2). However, such repetitive elements are widely present in all mammalian genomes. For example, up to 50% of the human genome is derived from repetitive elements (Smit 1996; Kazazian 1998).

Figure 23 illustrates the nucleic acid sequence (SEQ ID NO:7) of the Lama2/APPA transgene construct.

The Lama2 high level expression cassette (Laursen and Hjorth 1997) contains the enhancer region and the promoter of the Psp gene in the parotid gland. High expression was

shown to be dependent on regulatory elements between -11.5 kb and -6.5 kb and/or between +8.3 kb and +10.9 kb. Svendsen et al. (1998a) showed that a 1.5 kb sequence between -3.1 kb and -4.6 kb had properties of a parotid and sublingual specific enhancer and was designated as the PSP proximal enhancer. Furthermore, they showed that transgenes containing the PSP promoter and 5' flanking region located between -3.6 kb and -4.3 kb contained sequence information necessary to direct salivary gland specific expression.

Screening the transgene with RepeatMasker did not reveal the presence of any fulllength active autonomous elements. The repeats present were extensively modified by insertions and deletions. The blasts program was also used to compare the transgene 10 sequence translated in all reading frames against the National Center for Biotechnology Information (NCBI) protein sequence database (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul et al. 1990; Gish and States 1993; Terada and Nakanuma 1993). A region of DNA from 861 to 2180 was found that might code for parts of a protein with limited homology (38-58% identities) to the C-terminus of several human and mouse reverse transcriptases. 15 However, the region was extensively modified by mutations with multiple frame shifts and inversions, and probably represented remnants left from the reverse transcriptase gene of a LINE element. It is unlikely that it would be active, due to extensive modifications in the amino acid sequence such that only 18% of the full reverse transcriptase sequence was present and the highly conserved amino acid motif (Y/FXDD) was absent from the sequence 20 (Xiong and Eickbush 1990). The complete sequence was also scanned for the presence of open reading frames (ORFs) that code for proteins using the program GENSCAN (http://CCR-081.mit.edu/GENSCAN.html) (Burge and Karlin 1997). Only one gene was found and it corresponded to the APPA phytase gene. GENSCAN unexpectedly predicted a different N-terminus for the phytase than would have been expected from the sequence. 25 However, that could have resulted from the lower accuracy of GENSCAN for detecting initiation sites (Burge and Karlin 1998).

3) Generation of Transgenic Mice Expressing a Constitutive Salivary Phytase

In the following description, a pair of founder mice, incorporating the phytase gene and a constitutive promoter, were prepared under contract by the University of Alabama. As will be discussed, these founders were used to produce offspring, which were then analyzed for the presence of the phytase gene by PCR and animals containing the gene were then tested constitutive salivary phytase production.

10

15

20

25

30

Two transgenic founder mice (a black male and a white female, 3-1) containing the phytase transgene were received from the NICHD Transgenic Mouse Development Facility at the University of Alabama. The black male was negative for salivary phytase, but the female, 3-1, exhibited a salivary phytase activity of 30 U/ml. Progeny produced by crossing the black male with 4 CD-1 females produced 9 out of 25 females and 13 out of 26 males that were PCR positive. All progeny were negative for salivary phytase. The female founder, 3-1, was out-crossed with a CD-1 male to produce 3 litters for a total of 35 offspring. Of the progeny from these matings one phytase positive G1 male was obtained. When the G1 male was outcrossed with 6 CD-1 females, of the 6 litters 20/34 males were PCR positive and salivary phytase positive and 21/28 females were PCR positive and salivary phytase positive (Table 3). The salivary phytase activity of different offspring from the same first generation (G1) male ranged from 1.3 to 71.2 U/ml. There was no significant difference in the phytase activities between male or female mice.

PCR assays for identification of the transgenic mice were carried out with an initial heating step at 95°C for 3 min, 40 cycles using 95°C for 30 sec, 54°C for 30 sec and 72°C for 1 min) using the following primers: APPA-UP2 and APPA-KPN (Figure 6).

The phytase assays were conducted as described above for the R15-PRP/APPA phytase expressed in cell culture.

4) Production of Transgenic Pigs Containing the Phytase Transgene Lama 2/APPA

Transgenic pigs were produced using Yorkshire and Yorkshire/Landrace cross gilts as the embryo donors and Yorkshire sows as the recipients. The experimental procedure used was similar to that described by Wall et al. (1985). The detailed procedure is described below. The Lama2/APPA construct with the APPA signal peptide was used as the transgene for microinjection.

Methodology for the generation of transgenic pigs

The following is a description of the preferred method of generating transgenic pigs according to the invention. However, it will be apparent to those skilled in the art that various other methods are also applicable.

a) Superovulation of prepuberal gilts and sows.

Selected Yorkshire or Yorkshire/Landrace cross gilts between 70 to 80 kg were superovulated by intramuscular injection of 2000 IU of pregnant mare's serum gonadotropin

(PMSG, Ayerst Veterinary Laboratories), followed by 700 IU human chorionic gonadotropin (HCG, Ayerst Veterinary Laboratories) 60 to 72 hours later, administered in the same manner. The gilts were artificially inseminated three times with a 16 hour interval between inseminations using semen from a high breeding index Yorkshire boar. Twenty-four hours after the last insemination, the gilts were slaughtered and the reproductive tract recovered.

b) Synchronization of estrus in recipients

Estrus was synchronized in experienced recipient sows as described for donor sows. Since synchronization and not superovulation was the goal, hormone levels were reduced to 500 IU for PSMG and 500 IU for HCG. PMSG was given the day the sow's litter was weaned, followed in 72 hours by HCG and surgery for embryo transfer was performed 54 hours thereafter.

c) Embryo collection

Reproductive tracts were collected at the abattoir, inserted into bags, sealed and the bags immersed in water at 39°C for transport to the laboratory. Recovery of the embryos and microinjection with the transgene was conducted in a laboratory maintained at 32 to 33°C. The oviducts were dissected from the tracts and flushed, using a syringe and a feeding tube, with 15 ml of pre-warmed HBECM-3 medium (Dobrinsky et al. 1996). The media was collected in a 100 mm Petri dish and placed in an incubator at 38.5°C with an atmosphere of 5% (vol/vol) of CO₂, 5% (vol/vol) O₂ and the balance N₂. After all tracts were flushed, embryos were individually collected from the flushed media using a polished transfer pipette. Embryos were rinsed twice in 3 ml volumes of pre-incubated BECM-3 and placed in 100 µl of pre-incubated BECM-3 under 3 ml of filter sterilized mineral oil until injected.

25

30

20

5

10

15

d) Pronuclear injection

Embryos from one gilt were collected and placed in one ml of pre-warmed HBECM-3 in a 1.5 ml centrifuge tube and centrifuged for 6 min at 14,000 x g (Wall et al. 1985). The embryos were then collected and placed in an injection dish with 40 μl of pre-warmed HBECM-3 covered with 2.5 ml of filter sterilized mineral oil. The pronucleus in each embryo was injected (Gordon et al. 1980) with three picolitres of Lama2/APPA DNA in solution at a concentration of 5 ng of DNA per μl in 10 mM Tris, pH 7.5, 0.1 mM EDTA. After injection, the embryos were placed in dishes containing 100 μl of pre-incubated

BECM-3 under 3 ml of filter sterilized mineral oil. After all embryos were injected, which took no more than 4 hours since collection of reproductive tracts, the embryos were transferred to 1.8 ml cryotube (Nunc) containing 1 ml of pre-warmed HBECM-3 and transported in an incubator at 38.5°C to the swine surgery.

5

10

15

e) Embryo transfer

Recipient sows were anesthetized by intravenous injection of 500 mg Brietol and anesthesia maintained by inhalation of 3% halothane with 4 litres per min of nitrous oxide and 4 litres per min oxygen. The oviducts were exposed through a laparotomy, just off the dorsal midline, and a catheter, containing 20 to 35 injected embryos and 3 to 6 untreated embryos, was passed into the infundibulum and down the oviduct to the isthmus and emptied. The oviduct was returned to the abdominal cavity and the incision closed.

f) Growth of pigs

New-born piglets were kept together until weaning. At that time males and females were separated and penned with non-transgenic same sex pigs of a similar age from other litters. The pigs are fed *ad libitum* starter rations until 25 kg wt, grower diet from 25 to 60 kg wt and finisher diet from 60 kg to market weight. Water is available *ad libitum*. Transgenic pigs 167-02, 282-02 and 282-04 were maintained on a low phytate ration until 85, 50, and 50 days of age, respectively, and then switched to the grower ration. All other transgenic pigs were given the standard high phosphorus diets.

The diets were provided as pelleted formulations during the weanling, grower and finishing phases are shown in Tables 4 and 5. The vitamin and mineral mixes included in the diets are shown in Tables 6 and 7.

25

30

20

PCR analysis

Tail segments from newborn piglets were collected and slices of each placed in 600 µl of 50 mM NaOH and heating at for 95°C for 15 minutes. The suspension was neutralized with 50 µl of 1 M Tris (pH 8.0) and insoluble materials removed by centrifugation for 5 min in a microcentrifuge. A 2 µl sample of each was used for PCR with primers APPA-UP2 and APPA-KPN.

The primers produce a 750 bp fragment if the transgene is present. As a positive control PIG-BGF and PIG-BGR primers were used to detect the porcine β-globin gene from

10

15

20

25

30

the same DNA preparation (Heneine and Switzer 1996). The PCR reaction was performed using the same conditions as described for detection of the phytase transgene. As a negative control genomic DNA from a non-transgenic pig was used in the PCR reaction, for a positive control this DNA was spiked with a known amount of transgene (1 gene copy/per genome).

When a positive signal was identified by PCR for pig 167-02 (Figure 3) another DNA preparation was made and two more pairs of PCR primers were used to test for gene integrity (Figure 4) APPA-MATURE with APPA-KPN, and APPA-MATURE with APPA-DOWN2

PCR conditions were similar to those described previously.

Extraction of DNA from blood for PCR analysis

The method for extraction of DNA from blood was based on a method described by Higuchi (1989) with some modifications. A 100 μl volume of whole blood was mixed with 200 μl of lysis buffer (10 mM Tris-HCl, 0.32 M sucrose, 5 mM MgCl₂, 1% (vol/vol) Triton X-100, pH 7.5.), mixed briefly and incubated on ice for 5 min. The sample was then centrifuged at 14,000 x G for 3 min, and the supernate discarded. The sediment was suspended in lysis buffer, mixed, incubated and centrifuged. This procedure was repeated 2 more times, or until no hemoglobin remained. The sediment was dissociated in 100 μl of 50 mM NaOH, mixed and heated at 100°C for 10 min. The contents were cooled, 10 μl of 1 M Tris-HCl (pH 8.5) added and mixed briefly. The sample was then centrifuged at 14,000 x g for 2 min and 2 μl of the supernate used for analysis by PCR.

The PCR reaction mixture with a total volume of 40 µl consisted of; 23.8 µl of distilled water, 4 µl of 10 X Gibco BRL PCR buffer, 1.2 µl of 50 mM MgCl₂, 0.8 µl of 10 mM dNTPs, 40 pmol of each of the forward and reverse primers in 8 µl, 2 µl of template DNA and 0.2 µl of Taq DNA polymerase (Gibco BRL, 5 U/µl). The amplification procedure was performed with an initial heating step at 95°C for 3 min followed by 40 cycles of 95°C for 30 sec, 54°C for 30 sec and 72°C for 60 sec.

The transgenic pigs were detected with primers for the APPA gene (APPA-KPN with APPA-UP2), and as a control PIG-BGF with PIG-BGR primers were used for detection of the porcine β -globin gene.

Saliva collection from pigs for phytase assays and weighing of pigs

Weanling pigs were sampled for salivary phytase by wiping under the tongue with a cotton tipped applicator, breaking the stick off and centrifuging the applicator tip in a 0.4 ml

microcentrifuge tube, with a hole in the bottom, contained within a 1.5 ml microcentrifuge tube. Grower and finishing pigs were sampled using 1.5 inch long #2 dental cotton absorbent rolls (Ash Temple Sundries Ltd, Don Mills, ON) attached to dental floss. These were centrifuged in 1.5 ml microcentrifuge tubes with holes in the bottom while contained in larger tubes. The saliva was collected from the larger tube and stored at -20°C until analyzed.

Saliva was collected and pigs were weighed at weekly intervals.

Analysis for phytase activity.

Saliva samples were either assayed directly or after dilution in 0.1 M acetate buffer pH 4.5. Phytase was assayed in 200 µl of 0.1 M sodium acetate buffer (pH 4.5) using sodium phytate (4 mM) as a substrate at 37°C. After 10 min of incubation the reaction was stopped by addition of 133 µl ammonium molybdate/ammonium vanadate/nitric acid mixture and the concentration of liberated inorganic phosphate determined at 405 nm (Engelen, van der Heeft, Randsdorp, and Smit 1994). This and all other assays were performed in triplicate. One unit (U) of enzyme activity was the amount of the enzyme releasing 1 µmol of inorganic phosphate per minute.

Assays for salivary phytase and for phytase in blood samples were conducted as previously described for saliva samples. A reagent blank with blood added at the same concentration as the samples assayed was subtracted from the sample readings.

20

25

.30

5

10

15

Collection of fecal materials and analysis for total phosphorus

Fresh feces were collected from each pig during the grower and finisher phases. Samples were placed in aluminum trays closed with a wax paper top and immediately frozen, and kept frozen until they were lyophilized for analysis. After lyophilization the samples were transferred to room conditions overnight to reach equilibrium in moisture content. The samples were separately ground with a mortar and pestle until homogenous and sealed in plastic containers until analyzed further. Dry matter content of samples was analyzed according to AOAC (Association of Official Analytical Chemists (AOAC) 1984) by heating 1 gram samples at 110°C for 4 hours and cooling in a desiccator prior to weighing. To analyze total phosphorus content, samples were heated at 550°C in a muffle furnace and 10 ml of 10 M HCl added and heated to boiling. The contents from each sample was quantitatively diluted to 250 ml with water and inorganic phosphorus content was measured by the method of Heinoen and Lahti (1981).

10

15

20

25

30

Purification of the E. coli produced phytase and pig salivary phytase

The APPA phytase was over expressed in *E. coli* strain BL21(DE3) and the EDTA lysozyme extract fraction purified on DEAE-Sepharose and Sephadex-G75 as described by Jia et al. (1998). The pig phytase was purified by chromatography on DEAE-Sepharose and the band of enzyme eluted with a sodium chloride gradient was further purified by Chromatofocusing using a pH gradient from pH 4.0 to 7.0.

SDS-PAGE analysis and Silver Staining

Sodium dodecylsulfate polyacrylamide gel electrophoresis was performed using a 10% gel as described by Laemmli (1970), except that protein in the sample buffer was heated at 70°C for 10 minutes. Samples were stained with silver as described by Nesterenko et al. (1994).

Preparation of a monoclonal antibody specific for the APPA encoded E. coli phytase

Monoclonal antibodies specific to the *E. coli APPA* encoded phytase were prepared according to the procedures of Galfrè and Milstein (1981). Briefly, two female Balb/c mice were immunized 7 times over a period of 59 days with a purified APPA enzyme preparation. Mouse spleens were harvested, and the cells therein fused with an NS-1 myeloma cell line (Kohler and Milstein, 1976). Fused cells were selected for their ability to grow in media containing hypoxanthine, aminopterin, and thymidine (HAT). Western blotting and Enzyme-Linked Immunosorbent Assays (ELISA) were used identify those clones capable of secreting an antibody into the culture medium that recognized epitopes on both the *E. coli* and pig derived APPA enzyme. Clones secreting a desirable antibody were subcloned twice to ensure a pure culture of antibody secreting hybridomas.

Production of Polyclonal Antibodies Against the Purified E. coli derived APPA Phytase

Antibodies were prepared in two New Zealand White Rabbits by two intramuscular injections at different sites in the thigh of 50 µg of purified Escherichia coli derived APPA phytase in 0.5 ml of a 1:1 mixture of phosphate-buffered saline (PBS) and Freund's Complete Adjuvant. This was followed by repeat injections of 20 µg each of phytase in a 1:1 mixture of PBS and Freund's Incomplete Adjuvant on days 4, 19, 25, and 39. Blood was collected via heart puncture on day 42. The serum was separated from the cell fraction and used as the

source of antibodies. The basic procedures for antibody production are described in Harlow and Lane (1988).

Western blotting

Western blotting was performed as described by Towbin et al. (Towbin et al. 1979).

Deglycosylation of pig phytase was done according to protocols, Roche Molecular

Biochemicals, with following modifications. Protein in 50 mM Tris (pH 8.0), 1 mM EDTA,
1% SDS, 1% 2-mercaptoethanol was denaturated by heating at 95° C for 3 min. Than protein
was precipitated with chloroform-methanol method (Wessel and Flugge 1984) and
resuspended at 100 μg/mL in 20 mM Sodium Phosphate (pH 7.2) with 1% Triton X-100.

Complete deglycosylation of 5 μg in 50 μL phytase was carried out overnight at 37°C using
1 unit (U) N-glycosidase F, 1.2 mU O- glycosidase and 1 mU neuraminidase (Boehringer

Mannheim GmbH). After incubation 0.5 μg of protein was run on the SDS gel.

15 Staining of glycoproteins

20

25

.30

This staining was done using DIG Glycan Detection Kit (Boehringer Mannheim) according to manufacture instructions (O'Shannessy et al. 1987).

Statistics on the generation of transgenic pigs

The statistics on embryos recovered, microinjected and transferred into donor sows is shown in Table 8. A total of 4147 embryos injected with the transgene and 675 untreated embryos were introduced into 140 recipient sows with an average of 30 injected embryos and 5 uninjected embryos. All offspring were tested for the presence of the transgene in tissue biopsy, in blood by PCR analysis, and by an assay for phytase activity in the saliva.

Table 9 lists the transgenic pigs that were produced, their birth dates, sex and salivary phytase levels. There were 31 pigs transgenic for the phytase gene out of 203 live piglets born from embryos microinjected. These were detected by the presence of the gene in blood samples using the standard primer set, APPA-UP2 and APPA -KPN, but only 14 were detected by analysis of tail DNA preparations using the standard primer set. When the negative samples were reanalyzed using the primer set LAMA-UP1 and APPA-down4 (Figure 8) a further 8 tail DNA samples were found to be positive. Purification of the tail biopsy DNA probably would have led to all being PCR positive for the phytase transgene.

15

25

30

Characteristics of the phytase transgene in transgenic pig 167-02

The application of PCR to detection of transgenic pigs is exemplified by analysis of litter 167 in which one of 7 piglets tested, including one that was stillborn and one that was crushed by the sow after birth, one live piglet designated 167-02 was identified as positive for the APPA gene by generation of a PCR product (Lane 2) of approximately 750 bps from the tail chromosomal DNA (Figure 7). No rearrangements of the APPA gene were detected as documented by the positive PCR results using primers directed to the 3' region (lane 2) the whole gene (lane 3) and the 5' region (lane 4) of the APPA gene (Figure 8).

10 Salivary phytase and weight gain during growth of transgenic and non-transgenic penmates.

Data on salivary phytase activity and weight gain are shown for five transgenic pigs and for weight gains of their non-transgenic penmates in Figures 9, 10, 11, 12 and 13. The phytase activity in the saliva varied substantially from one sampling time to the next. This variability was attributed to a combination of environmental factors including whether the animal had just consumed food or water, and regulation of parotid and saliva secretion in relation to food and water consumption. The weight gains during growth of the five transgenic pigs was within the range of the weight gains of the normal non-transgenic pigs.

With the exception of 167-02 the growth rate of the transgenic pigs was similar to that of the non-transgenic litter mates.

20 Phosphorus content in the fecal materials from transgenic and non-transgenic pigs.

The phosphorus content of fresh fecal samples from three of the transgenic founder pigs, 167-02, 282-02, 282-04, 405-02 and 421-06 receiving weaning, grower or finisher ration is shown in Table 9. The phosphorus content of the feces of the transgenic pigs ranged from 1.59 to 2.26% while that of the non-transgenic penmates ranged from 1.61 to 2.76%. The reduction in fecal phosphorus ranged from a maximum of 26% to a minimum of 8%. In most cases the differences were at the 99% level of significance. The ages of the pigs at the time of fecal sampling and the corresponding phytase activities are shown in Figures 9, 10, 11, 12 & 13. The rations fed contained a supplement of readily available phosphorus suitable for maximizing growth of non-transgenic pigs. Since the reduction in fecal phosphorus is measured in transgenic pigs receiving a diet high in mineral phosphorus it is very likely that the fecal phosphorus would be substantially lower if the diet lacked mineral phosphorus. Under these conditions the phosphorus released from phytate would provide a substantial

10

15

20

25

30

proportion of the dietary phosphorus and little would reach the large intestine and be excreted in the feces.

Transmission of the phytase transgene (to be completed)

When semen from the transgenic boar 167-02 was used to inseminate four Yorkshire gilts all four sows had litters in which 4 out of 8, 2 out of 9, 7 out of 8 and 2 out of 5 of the piglets were transgenic for the phytase gene (Table 11). The PCR data for litter 154 that documents the presence of the transgene is shown in Figure 14. All pigs containing the gene exhibited phytase activity in the saliva, and it ranged from 341 to 10,077 units per ml. Half of the transgenic piglets had salivary phytase activities of greater than 2000 units per ml. The specific activity of the phytase in the saliva ranged from 39 U/mg protein to a high of 706 units/mg protein.

This data documents that the gene was transferred and that the level of phytase expression observed in the founder was preserved in the first generation of pigs. Both male and female pigs at 11 days of age exhibited high phytase activity.

Characteristics of the phytase enzyme synthesized in the salivary glands of the pig

The phytase enzyme was purified to homogeneity from E. coli and from saliva collected from transgenic pig 167-02. Silver stains of the purified enzymes after SDS-PAGE are shown in Figure. 15. The E. coli derived enzyme has a molecular mass of approximately 45 kDa while that produced by the pig was about 55 kDa. The enzymes were also electrophoresed as before, transferred to nitrocellulose and stained for glycoproteins. The second part of Figure 15 shows that the pig APPA protein is glycosylated. Figure 15B shows that treatment of the pig phytase with deglycosylation enzymes changes the size of the phytase from 60 kDa to 45 kDa, an observation that confirms the glycosylated nature of the recombinant phytase produced in the saliva of the pig.

The data in Figure 16 shows that the pig phytase is homologous with the *E. Coli* enzyme despite their difference in size.

The purified pig phytase had K_m and V_{max} values of 0.33 mM and 624 units per mg of protein, respectively. Golovan et al. (2000) previously reported the K_m and V_{max} for the E. coli enzyme to be 0.63 mM and 2325 units per mg of protein. Thus the salivary phytase exhibits approximately 25% of the activity of the E. coli enzyme. This reduction in activity may be due to glycosylation that either modifies the catalytic site of the enzyme or otherwise leads to the formation of an enzyme with lower catalytic activity.

10

15

25

30

The latter finding of the production of a glycosylated protein suggests a method of producing such proteins using transgenic animals. Currently, although recombinant methods are available for producing proteins in host cells, it is often found that the mature peptide lacks the glycosylation normally associated with proteins produced by higher life forms. Insulin is an example of such protein. The findings of this study suggest that one means of

Insulin is an example of such protein. The findings of this study suggest that one means of producing the desired glycoproteins would be to generate transgenic animals such as the pig, that have been transformed, by known methods or the method described above, with a gene encoding the desired protein. When expressed by such animal, the subject protein would be produced and would undergo post-translational processing in the cell including the step of glycosylation. Thus, the invention contemplates a general method of producing such glycosylated proteins. Further, the invention contemplates a method of producing glycosylated proteins through the expression in and isolation from the saliva of an animal that has been transformed with a gene encoding such protein, and wherein such gene is operably linked to a saliva protein promoter or enhancer.

Various methods are known in the art for the collection of glycoproteins from the parotid gland of the pig for various applications. For example, surgical techniques have been published by Denny et al. (1972) for the collection of secretions from the parotid gland and submandibular salivary ducts.

20 Test kit for detection of the APPA phytase protein in pigs

The monoclonal antibodies produced against the APPA phytase expressed in *E. coli* reacted with the APPA phytases produced in the saliva of transgenic mice and pigs (Figure 17). Immunological detection of phytase in saliva provides definitive proof that the phytase secreted in transgenic pig saliva is a product of the *APPA* gene expressed in the pig salivary gland. This serves as a reliable method to document phytase production in transgenic pigs.

A further test would also be obtainable using the polyclonal antibodies discussed above.

The DNA sequence encoding phytase may be obtained from a variety of sources such as microbial, plant or animal sources. Preferably, the DNA sequence is obtained from a microbial source such as bacteria. Most preferred DNA sequences are obtained from Escherichia coli.

The cloning of a gene or a cDNA encoding a phytase protein may be achieved using various methods. One method is by purification of the phytase protein, subsequent

10

15

20

25

determination of the N-terminal and several internal amino acid sequences and screening of a genomic or cDNA library of the organism producing the phytase using oligonucleotide probes based on the amino acid sequences. If at least a partial sequence of the gene is known, this information may be used to clone the corresponding cDNA using, for instance, the polymerase chain reaction (PCR) (PCR Technology: Principles and Applications for DNA Amplification, (1989) H. A. Ehrlich, ed., Stockton Press, New York; the contents of which are incorporated herein by reference). It will be evident to those skilled in the art that the cloned phytase gene described above may be used in heterologous hybridization experiments, directed to the isolation of phytase encoding genes from other microorganisms.

The DNAs encoding phytase or individual fragments or modified proteins thereof can be fused, in proper reading frame, with appropriate regulatory signals as described in detail below, to produce a genetic construct that is then amplified, for example, by preparation in a bacterial (e.g., E. coli) plasmid vector according to conventional methods. Such methods are described in, for example, Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Press 1989), the contents of which are incorporated herein by reference. The amplified construct is thereafter excised from the vector and purified for use in producing transgenic animals.

The desired protein may also be produced as a fusion protein containing another protein. For example, the desired recombinant protein of this invention may be produced as part of a larger recombinant protein in order to stabilize the desired protein. Useful modifications within this context include, but are not limited to, those that alter post-translational modifications, size or active site, or that fuse the protein or portions thereof to another protein. Such modifications can be introduced into the protein by techniques well known in this art, such as by synthesizing modified genes by ligation of overlapping oligonucleotides or introducing mutations into the cloned genes by, for example, oligonucleotide-mediated mutagenesis.

The cloned phytase gene may be used as starting materials for the construction of improved phytases. Improved phytases are phytases, altered by mutagenesis techniques (e.g. site-directed mutagenesis, or directed evolution), which have properties that differ from those of wild-type phytases (Kuchner and Arnold 1997). For example, the temperature or pH optimum, specific activity, temperature or protease resistance may be altered so as to be better suited for a particular application.

A choice of expression in cellular compartments (such as cytosol, endoplasmic reticulum) or extracellular expression can be used in the present invention, depending on the

10

15

20

25

30

biophysical and biochemical properties of the phytase. Such properties include, but are not limited to pH sensitivity, sensitivity to proteases, and sensitivity to the ionic strength of the preferred compartment. The DNA sequence encoding the enzyme of interest should be modified in such a way that the enzyme can exert its action at the desired location in the cell. To achieve extracellular expression of the phytase, the expression construct of the present invention utilizes a bacterial signal sequence. Although signal sequences that are homologous (native) to the animal host species are preferred, heterologous signal sequences, i.e. those originating from other animal species or of microbial origin, may be used as well. Such signal sequences are known to those skilled in the art.

All parts of the relevant DNA constructs (promoters, regulatory, secretory, stabilizing, targeting, or termination sequences) of the present invention may be modified, if desired, to affect their control characteristics using methods known to those skilled in the art. The cisacting regulatory regions useful in the invention include the promoter that drives expression of the phytase gene. Highly preferred are promoters that are specifically active in salivary gland cells. Among such promoters, highly preferred are mouse parotid secretory protein (PSP) promoter, rat proline-rich protein (PRP) promoter, human salivary amylase promoter, mouse mammary tumor virus promoter (Samuelson 1996). Among the useful sequences that regulate transcription, in addition to the promoters discussed above, are enhancers, splice signals, transcription termination signals, and polyadenylation sites. Particularly useful in this regard are those that increase the efficiency of the transcription of the genes for phytase in the salivary gland or other cells of the transgenic animals listed above. Preferred are transcription regulatory sequences for proteins highly expressed in the salivary gland cells. Introns could be introduced to increase levels of expression. Such introns include the synthetic intron SIS, SV40 small t antigen intron and others (Whitelaw et al. 1991; Petitclerc et al. 1995).

Preferably, the expression system or construct of this invention also includes a 3' untranslated region downstream of the DNA sequence encoding the desired recombinant protein, or the salivary protein gene used for regulation. This region apparently stabilizes the RNA transcript of the expression system and thus increases the yield of the desired protein. Among the 3' untranslated regions useful in this regard are sequences that provide a polyA signal. Such sequences may be derived, e.g., from the SV 40 small t antigen late polyadenylation signal, synthetic polyadenylation signal or other 3' untranslated sequences well known in this art (Carswell and Alwine 1989; Levitt et al. 1989). Preferably, the 3' untranslated region is derived from a salivary-specific protein. The stabilizing effect of this

10

15

20

25

30

region's polyA transcript is important in stabilizing the mRNA of the expression sequence. Further, the addition of locus control regions (LCRs), matrix attachment regions (MAR) and scaffold attachment regions (SARs) would allow position-independent, copy number dependent expression of the transgene with either homologous or heterologous promoters (Taboit-Dameron et al. 1999; Geyer 1997). Co-integration of an actively expressed gene with the transgene was also shown to increase expression levels of a poorly expressed transgene (Clark et al. 1993). Also important in increasing the efficiency of expression of phytase is a strong translation initiation site (Kozak 1987). Likewise, sequences that regulate the post-translational modification of phytase may be useful in the invention.

The term "animal" as used herein denotes all animals except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages.

A "transgenic" animal is any animal containing cells that bear genetic information received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by microinjection or infection with a recombinant virus. "Transgenic" in the present context does not encompass classical crossbreeding or in vitro fertilization, but rather denotes animals in which one or more cells receive a recombinant DNA molecule. Although it is highly preferred that this molecule be integrated within the animal's chromosomes, the invention also encompasses the use of extrachromosomally replicating DNA sequences, such as might be engineered into yeast artificial chromosomes. The information to be introduced into the animal may be foreign to the species of the animal to which the recipient belongs (i.e., "heterologous"), or the information may be foreign only to the particular individual recipient, or genetic information already possessed by the recipient. In the last case, the introduced gene may be expressed in a manner different than the native gene.

As indicated above, the transgenic animals of this invention are other than human. Farm animals (pigs, goats, sheep, cows, horses, rabbits and the like), rodents (such as mice and rats), domestic pets (eg. cats and dogs), fish and poultry (eg. chickens) are included in the scope of this invention. It is highly preferred that a transgenic animal of the present invention be produced by introducing into single cell embryos appropriate polynucleotides that encode phytase, or fragments or modified products thereof, in a manner such that these polynucleotides are stably integrated into the DNA of germ line cells of the mature animal, and are inherited in normal mendelian fashion. Advances in technologies for embryo micromanipulation now permit introduction of heterologous DNA into fertilized mammalian ova. For instance, totipotent or pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other

15

20

means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In one preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a most preferred method, however, the appropriate DNAs are co-injected into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals. Such techniques are well known (see reviews of standard laboratory procedures for microinjection of heterologous DNAs into mammalian fertilized ova, including Hogan et al., Manipulating The Mouse Embryo, (Cold Spring Harbor Press 1986); Krimpenfort et al., Bio/Technology 9:844 (1991); Palmiter et al., Cell, 41: 343 (1985); Kraemer et al., Genetic Manipulation Of The Early Mammalian Embryo, (Cold Spring Harbor Laboratory Press 1985); Hammer et al., Nature, 315: 680 (1985); Wagner et al., U.S. Pat. No. 5,175,385; Krimpenfort et al., U.S. Pat. No. 5,175,384, the respective contents of which are incorporated herein by reference).

For a person skilled in art, it will also be clear that the present invention provides for other proteins to be expressed in the salivary gland of the pig. Such proteins may be secreted into saliva to improve digestion and decrease pollution potential (for example, endoglucanases), or specifically targeted for secretion into blood and have effects on the growth and health of the animal (such as growth hormone).

Phytase activity may be measured via a number of assays, the choice of which is not critical to the present invention. For example, the phytase enzyme activity of the transgenic animal tissue may be tested with an ELISA-assay, Western blotting or direct enzyme assays using calorimetric techniques or gel assay system.

The examples included herein are provided so as to give those of ordinary skill in the art a complete disclosure and description of how to make and use the invention and are not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperature, pH, etc.) but some experimental errors and deviation should be accounted for. Unless indicated otherwise, temperature is in degrees Centigrade and pressure is at or near atmospheric.

30

25

Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention as outlined in the claims appended hereto.

Table 1. Secretion of phytase in the saliva of transgenic mice containing the R15-PRP/APPA transgene and non-transgenic mice induced with isoproterenol and pilocarpine.

Founder	Mice	PCR	Gender	Generation	Transgene	Phytase activity
A0m	4bfr (+)	positive	F	1	APPA+intron	39.73
A0m	2brm(+)	positive	M	1	APPA+intron	24.29
A0m	2brm(+)	positive	M	2	APPA+intron	14.42
A0m	5brf(+)	positive	F	2	APPA+intron	7.36
A0m	1brm(-)	negative	M	1	APPA+intron	0.00
Alf	9brf(+)	positive	F	1	APPA+intron	0.08
Alf	11w f(+)	positive	F	1	APPA+intron	0.07
A1f	5brm(+)	positive	M	1	APPA+intron	0.03
A1f	10wf(-)	negative	F	1	APPA+intron	0.02
A20f	1brm(+)	positive	·M	1	APPA+intron	0.53
A20f	5brf(+)	positive	F	1	APPA+intron	0.12
A20f	4brf (-)	negative	F	1	APPA+intron	0.03
A2m	13wf(+)	positive	F	1	APPA+intron	87.70
B0m	5brf (+)	positive	F	1	APPA+intron	0.95
B0m	3brm(+)	positive	M	1	APPA+intron	0.73
B0m	6wf (-)	negative	F	1	APPA+intron	0.00
B0f	3wf (+)	positive	F	2	APPA	252.43
B0m-intr	9wf(+)	positive	F	1	APPA .	546.74
W0m	8wf(+)	positive	F	1	APPA	60.42
W30m	lwm(+)	positive	M	2	APPA	41.91
W30m	11w f(+)	positive	F	1	APPA	43.44
W30m	4wm(-)	negative	M	1	APPA	0.02
W30m	10wf(-)	negative	F	1	APPA	0.02

Table 2. Repeat sequences found in the Lama2-APPA constitution.

Start	End	DNA	Repeat	Class/family	Substitu-	Deletions	Insertions.
Start	الملاد	strand	Ropean	J	tions % of	% of	% of
1	1	Judic			consensus	consensus	consensus
765	927	+	LIMI	LINE/L1	25	4.2	6.7
928	965	+	(CA)n	Simple repeat	0	0	0
966	1020	+	LIMI	LINE/LI	25	4.2	6.7
1021	1156	+	B1 MM	SINE/Alu	15.4	0	0
1159	1231	+	CAAAC)n	Simple repeat	1.4	0	0
1232	1385	+	LIMI	LINE/L1	25	4.2	6.7
1652	2308	С	Ll	LINE/L1	28.5	11.9	1.7
2334	2406	С	MIR	SINE/MIR	27.4	4.1	0
2415	3266	+	RMER13A	LTR	17.7	4	6.1
6016	6127	С	L1MA9	LINE/L1	25.5	2	1
6831	7007	+	CT-rich	Low complexity	30.5	1.7	3.4
7299	7510	С	B3	SINE/B2	27.8	7.5	1.4
7718	7746	+	(TCTCTG)n	Simple repeat	6.9	0	0
8499	8581	С	MIR	SINE/MIR	24.1	12.1	3.6
9010	9603	+	Lx4	LINE/L1	21.7	6.4	0.2
10465	10519	+	(TG)n	Simple repeat	5.5	1.8	0
11235	11287	С	MER5A	DNA/MER1 type	28.3	0	1.9
12372	12537	С	L1MA4A	LINE/LI	28.3	5.4	0
14240	14388		B1_MM	SINE/Alu	4	0	1.3
14869	14945	C	MIR	SINE/MIR	36.4	1.3	0
16391	16540		ORR1D	LTR/MaLR	29.3	0	6
16774	17214		RMER4	LTR	21.3	10	11.8
17229			LI MM	LINE/L1	15.3	0	0.8

Table 3. Salivary phytase activities of G2 mice from the founder female 3-1 generated using the construct Lama2-APPA. The mice were between 21 and 30 days of age.

male mouse #	Phytase (U/ml)	female mouse #	Phytase (U/ml)
5	28.3	1	9.0
6	2.5	2	29.9
8	6.6	4	8.0
9	44.7	5	43.0
10	12.7	6	26.9
12	28.3	8	1.9
15	28.1	9	66.3
18	71.2	10	19.9
19	19.5	11	61.3
20	15.7	12	36.4
21	20.9	13	18.0
22	4.1	17	38.9
24	13.0	18	18.5
26	53.4	19	27.0
28	20.4	23	6.5
29	34.1	24	16.1
30	11.1	25	9.4
32	3.1	26	14.8
33	51.7	27	1.3
. 34	19.0	28	8.2

Table 4. Composition and nutrient levels of Phase II starter diet and low phytate starter diets fed to weanling pigs between 5-10 kg.

Ingredients		utrient Levels ¹
	Phase II Starter Diet	Low Phytate Starter Diet
Corn	33.15	25.44
Barley	8.00	8.00
Wheat	20.00	40.00
Soybean meal	21.00	8.00
Fish meal	5.00	5.00
Meat and bone meal	•	1.00
Whey	8.00	8.00
Fat	2.00	2.00
Lysine-HCl	0.10	0.28
Dicalcium phosphate	1.10	-
CaCO ₃	0.90	1.10
Iodized salt	0.30	0.30
Vitamin premix ¹	0.250	0.55
Mineral premix	0.10	0.10
Lincommix 44	0.10	0.10
Total (kg)	100.00	100.00
Total (kg)		
Calculated nutritive values		
DE (kcal/g)	3.44	3.36
CP (%)	19.46	18.62
Ca (%)	1.00	0.94
Total P (%)	0.74	0.66
Ca/P	1.35:1	1.42:1
Total AA contents (%)		
Arginine	1.16	1.17
Histidine	0.50	0.48
Isoleucine	0.81	0.77
Leucine	1.58	1.54
Lysine	1.17	1.06
Methionine	0.34	0.29
Cysteine	0.34	0.34
Methionine+Cysteine	0.68	0.63
Phenylalanine	0.90	0.90
Tyrosine	0.65	0.65
Threonine	0.75	0.68
Tryptophan	0.23	0.23
Valine	0.91	0.86

¹Minerals and vitamins meet or exceed levels recommended by NRC (1998).

Table 5. Composition and nutrient levels of grower and finisher diets.

Ingredients	Diet/Nu	trient Levels
	Grower Diet	Finishing Diet
	For pigs 20 to 50 kg	For pigs 50 to 120 kg
Com	51.78	40.00
Barley	8.10	23.03
Wheat	20.00	23.00
Soybean meal	16.00	13.00
Fat	1.00	1.00
Lysine-HCl	0.12	0.12
Dicalcium phosphate	1.20	1.00
CaCO ₃	1.15	. 1.15
Iodized salt	0.50	0.50
Vitamin premix ¹	0.15	0.15
Mineral premix	0.10	0.10
Total (kg)	100.00	100.05
Calculated nutritive values		
DE (kcal/g)	3.39	3.33
CP (%)	14.76	14.17
Ca (%)	0.79	0.74
Total P (%)	0.57	0.53
Ca/P	1.39:1	1.39:1
Total AA contents (%)		
Arginine	0.86	0.80
Histidine	0.38	0.36
Isoleucine	0.58	0.55
Leucine	1.28	1.18
Lysine	0.78	0.73
Methionine	0.24	0.23
Cysteine	0.29	0.29
Methionine+Cysteine	0.53	0.52
Phenylalanine	0.70	0.68
Tyrosine	0.50	0.46
Threonine	0.52	0.49
Tryptophan	0.17	0.16
Valine	0.68	0.65

¹Minerals and vitamins meet or exceed levels recommended by NRC (1998).

Table 6. Vitamin premix composition¹

Nutrient	Amount per 5 kg of premix	
Wheat midds	3.867 kg	
Vitamin A	10 million IU	
Vitamin D	1 million IU	
Vitamin E	40 thousand IU	
Menadione	2.5 g	
Pantothenic acid	15 g	
Riboflavin	5 g .	
Folic acid	2 g	
Niacin	25 g	
Thiamin	1.5 g	
Pyridoxine	1.5 g	
Vitamin B ₁₂	25 mg	
Biotin	· 200 mg	
Choline	500 g	

From Hoffman-LaRoche Limited, P.O. Box 877, Cambridge, ON. N1R5X9

Table 7. Composition of the mineral premix^{1,2}

Mineral component	Amount (%)	
Limestone	43.3	
Copper sulfate (25%)	6.0	
Ferrous sulfate (30%)	33.4	
Zinc oxide (72%)	13.9	
Manganous oxide (56%)	3.4	

Mineral premix prepared at Arkell

²Dicalcium phosphate contained 18.5% calcium and 20.5% of phosphate and normally is added at a level of 1.2% to the pig grower diet, 1.0% to the finisher diet and 1.5% to the nursing sow diet.

<u>Table 8. Statistics on embryo recovery and the introduction of embryos</u> containing the transgene into recipient sows.

Treatment Gilts used for embryo recovery: Yorkshire Yorkshire x Landrace cross Duroc Total Recipient sows! Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos Total	Number
Yorkshire Yorkshire x Landrace cross Duroc Total Recipient sows Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	
Yorkshire x Landrace cross Duroc Total Recipient sows Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	279
Duroc Total Recipient sows Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	168
Total Recipient sows ¹ Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	12
Recipient sows ¹ Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	459
Embryos transferred to recipients: Embryos microinjected with the transgene Uninjected carrier embryos	74
Embryos microinjected with the transgene Uninjected carrier embryos	17
Uninjected carrier embryos	4147
	675
Total	4543
Total number of embryo transfers	140

Sows were used for up to three farrowings of potentially transgenic pigs. Sows were inseminated with Yorkshire semen from a high breeding value boars.

Table 9. Transgenic pigs containing a salivary phytase gene generated by microinjections of single cell zygotes using the Lama?-APPA transgene

	le cell zygotes	using the Lama2-APPA			
ID # of	Birth Date	Presence of	Sex	Salivary phytase	Zygote source⁴
pig ¹		Transgene ² Tail/Blood		(U/ml) ³	
167-02	Apr 14/99	+/+	Boar	6,000	Yorkshire
282-02	Jun 14/99	+/+	Boar	618	Yorkshire
282-04	Jun 14/99	+/+	Boar	1,349	Yorkshire
405-02	Aug 14/99	+/+	Gilt	339	York/Landrace
421-02	Aug 24/99	-/+	Gilt	0.8	York/Landrace
421-04	Aug24/99	-/ +	Gilt	2.2	York/Landrace
421-06	Aug 24/99	+/+	Boar	97	York/Landrace
448-01	Sep 03/99	+/+	Gilt	0	York/Landrace
491-01	Sep 25/99	+/+	Gilt	2.3	York/Landrace
491-02	Sep 25/99	+/+	Gilt	0	York/Landrace
491-03	Sep 25/99	+/+	Gilt	0.3	York/Landrace
491-05	Sep 25/99	+/+	Boar	0	York/Landrace
496-05	Sep 26/99	+/+	Boar	0	York/Landrace
500-03	Sep 28/99	+/+	Boar	136	York/Landrace
510-01	Sep 28/99	+/+	Boar	0.2	York/
559-05	Nov 01/99	+*/+	Boar	>418	York/Landrace
560-04	Nov 02/99	+*/+	Boar	5	Yorkshire
594-03	Nov 18/99	+/+	Gilt	• 2.3	Yorkshire
613-02	Nov 27/99	-/ +	Gilt	0.5	York/Landrace
613-03	Nov 27/99	<i>-</i> /+	Gilt	0.3	York/Landrace
647-01	Dec 13/99	<i>-/</i> +	Gilt	0.5	York/Landrace
647-03	Dec 13/99	+*/+	Gilt	16.3	York/Landrace
647-04	Dec 13/99	- */ +	Gilt	0.5	York/Landrace
647-08	Dec 13/99	-*/+	Boar	0.4	York/Landrace
647-09	Dec 13/99	+*/+	Boar	1.92	York/Landrace
668-01	Dec 17/99	+*/+	Gilt	489	Yorkshire
671-02	Dec 19/99	+*/+ ·	Boar	6.9	York/Landrace
671-04	Dec 19/99	+*/+	Boar	325	York/Landrace
675-03	Dec 21/99	-*/ +	Gilt	2.1	York/Landrace
675-04	Dec 21/99	+*/+	Boar	42.6	York/Landrace
675-06	Dec 21/99	_*/+	Boar	5.0	York/Landrace

¹The number preceeding the dash represents the litter number and the number following the dash is the pig number within the litter.

²All PCR assays were conducted with the primer APPA-up2-APPA-Kpn. Assays indicated with a star gave a negative result with the primer pair. However these samples gave a positive result for the primer set APPA-d4-Lama-up1. Samples 613-02 and 613-03 were negative with the latter primer set.

³Saliva was sampled and assayed for phytase 2 to 4 days after birth of the piglets.

⁴Zygotes used for microinjection were collected from superovulated Yorkshire or Yorkshire-Landrace cross gilts.

Table 10. Phosphorus content of feces collected from pigs producing a salivary phytase and non-transgenic pen-mates¹. The data was subjected to a T-test analysis and the data recorded below.

	Mean Fecal	SE	Relative reduction	t	t (1%)
	Phosphorus		in fecal	Ť	()
<u> </u>	(%)		phosphorus (%)		
1. 167-02 Grower Diet (122 days):	1.59		24.47		
Non-transgenic (n=4)	2.11	0.0604669		8.517	4.6
2. 167-02 Finisher Diet (154	1.97		16.97		
days):					i
Non-transgenic (n=4)	2.37	0.0240767		16.717	4.6
3. 282-02 Grower Diet (93 days):	1.85		12.90		
Non-transgenic (n=5)	2.124	0.022231964		12.324	4.03
4. 282-02 Finisher Diet (145	1.76		16.03		
days):		ı			
Non-transgenic (n=5)	2.096	0.099153384		3.389	4.03 ²
5. 282-04 Grower Diet (93 days):	1.95		8.19		
Non-transgenic (n=5)	2.124	0.022231964		7.827	4.03
6. 282-04 Finisher Diet (145	1.56		25.57		
days):					
Non-transgenic (n=5)	2.096	0.099153384		5.406	4.03
7. 421-06 Starter II Diet (40	1.17		27.47		
days):					
Non-transgenic (n=5)	1.612	0.086155741		5.140	4.03
8. 421-06 Start III Diet (48 days):	1.57		18.01		
Non-transgenic (n=5)	1.915	0.102884789		3.351	4.03
9. 421-06 Grower Diet (81 days):	2.00		13.28		
Non-transgenic (n=5)	2.310	0.151658823		2.022	4.03
10. 421-06 Finisher Diet (136	1.71		21.20		
days):					
Non-transgenic (n=5)	2.173	0.053023237		8.687	4.03
11. 405-02 Starter II Diet (40	1.81		26.97		l
days):				ļ	
Non-transgenic (n=5)	2.482	0.173625623		3.856	4.03
12. 405-02 Starter III Diet (48	1.54		36.58		1
days):	ļ				
Non transgenic (n=4)	2.430	0.104642248		8.496	4.6
13. 405-02 Grower Diet (80 days):			18.19		
Non-transgenic (n=4)	2.763	0.124724697		4.029	4.6
14. 405-02 Finisher Diet (136	2.26		13.24		
days):	<u> </u>			<u> </u>	<u> </u>
Non-transgenic (n=4)	2.605	0.217198066	المناع		4.6

¹Fresh fecal samples were collected on 3 different days was freeze-dried and then dried to constant weight at 110°C for 24 h, and analyzed for total phosphorus.

²At the 5% level of confidence t=2.57.

Table 11. Phytase activities of the first generation (G1) transgenic offspring obtained by the crossing the phytase positive boar 167-02 with non-transgenic Yorkshire gilts¹

ID# of pig	Birth Date	Sex	Salivary phytase (U/ml)	Specific Activity U/mg protein
151-01	Mar 16/00	F	1193	126
151-02	ч	F	736	63.3
151-05	"	M	710	109
151-07	и	М	8019	315
152-04	"	M	10077	364
152-09	"	M	3054	200
154-01	Mar 19/00	F	2472	256
154-03	44	F	6425	706
154-04	"	F	n.d.	n.d.
154-05	44	M	2767	213
154-06	"	M	341	39
154-07	"	M	4029	142
154-08	46	M	1184	47.4
159-03	Mar 20/00	F	1563	116
159-04	и	M	2285	201

The number of males and females (M/F) in each litter were 5/3, 7/2, 5/4, and 2/3 for litter numbers 151, 152, 154 and 159, respectively. Saliva was collected from the piglets on day 11.

Table 12. Primers used for construction and detection of transgenic constructs.

Name	Start-End ¹	Forward/	
		Reverse	
Primers use	d in R15/APPA+	intron and R1	5/APPA construction
APPA-		R	TCGGCGCTCACCTTGAGTTC
DOWN2			
APPA-		F	CCGTTTAAAGCCATCTTAATCCCAT
DRA			
APPA-		R	GTCCCGGGTATGCGTGCTTCATTC
SMA			
CAT-ATG		R	CCATGGTGGCGGCTTTTAGCTTCCTTAGCTCCTGA
CAT-TAA		F	AGCGCTTGCAGTTTGTAAGGCAGTTATTG GTGCCC
CAT-UPI		F	TCG AGG AGC TTG GCG AGA TT
R15-UP1		F	TTTCGGGCCAATGTTGCTGT
KIJ-OL I	1	1.	
Primers use	d in SV40/APPA	tintron const	ruction
SV-HIND	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	F	CCCAAGCTTTACACTTTATGC
SV-XHO	-	R	GCCCTCGAGCCTCCTCACTACTTCT
31-7410	L		
Primers use	d in Lama2/API	A and Lama2	/PSP/APPA construction
APPA-	12635-12657	F	GGATCGATAAAAGCCGCCACCATGAA
CLA	12055 12057	1-	
APPA-	13307-13326	R	TCGGCGCTCACCTTGAGTTC
DOWN2	1		
APPA-	12751-12780	R	GCACGCACACCATGACGACTGACAATCAC
DOWN4			C
APPA-	13935-13959	R	CGGGTACCTTACAAACTGCAAGCGG
KPN			
APPA-	12719-12738	F	CAGAGTGAGCCGGAGCTGAA
MATURE			
APPA-	13210-13229	F	CGAACTGGAACGGGTGCTTA
UP2	<u></u>		
LAMA-	12615-12639	R	GCATCGATCTTTGGTTCTGACAAATGG
CLA			
LAMA-		R	TGACTCTGAGTTCCCAATGA
SIGNAL	<u> </u>		
LAMA-UP	12111-12130	F	GTGCTGCTCCAAGTTTGGTG
Primers fo	r detection of the	porcine β-glo	bin gene
PIG-BGF		F	GCAGATTCCCAAACCTTCGCAGAG
PIG-BGR	1	R	TCTGCCCAAGTCCTAAATGTGCGT

¹ The location of the primers shown for Lama2/APPA sequence.
The start and stop codons of APPA are indicated in bold letters, the optimal initiation sequence for translation is italicized, and the restriction sites for restriction enzymes are underlined.

Reference List

The following references have been referred to in the present application. The content of these references are incorporated herein by reference.

- 1. Abelson, P.H. 1999. A potential phosphate crisis [editorial]. Science 283: 2015.
- 2. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J.Mol.Biol. 215: 403-410.
- 3. Ann,D.K. and Carlson,D.M. 1985. The structure and organization of a proline-rich protein gene of a mouse multigene family. J.Biol.Chem. 260: 15863-15872.
- 4. Arkhipovets, A.I. 1956. Age characteristics of salivary secretion in young pigs. Journal of Physiology of USSR 42: 882-886.
- Association of Official Analytical Chemists (AOAC) 1984. Official methods of analysis of the Association of Official Analytical Chemists. Association of Official Analytical Chemists, Washington.
- 6. Ausbel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. 1992. Short protocols in molecular biology. John Wiley & Sons, New York.
- Barsh, G.S. and Epstein, C.J. 1989. Physical and genetic characterization of a 75-kilobase deletion associated with al, a recessive lethal allele at the mouse agouti locus.
 Genetics 121: 811-818.
- 8. Bennick, A. 1982. Salivary proline-rich proteins. Mol. Cell Biochem. 45: 83-99.
- Bitar, K. and Reinhold, J.G. 1972. Phytase and alkaline phosphatase activities in intestinal mucosae of rat, chicken, calf, and man. Biochim. Biophys. Acta 268: 442-452.
- 10. Burge, C. and Karlin, S. 1997. Prediction of complete gene structures in human genomic DNA. J.Mol.Biol 268: 78-94.

- Burge, C.B. and Karlin, S. 1998. Finding the genes in genomic DNA. 11. Curr.Opin.Struct.Biol. 8: 346-354.
- Carlson, D.M. 1993. Salivary proline-rich proteins: biochemistry, molecular biology, 12. and regulation of expression. Crit Rev. Oral Biol. Med. 4: 495-502.
- Carswell, S. and Alwine, J.C. 1989. Efficiency of utilization of the simian virus 40 late 13. polyadenylation site: effects of upstream sequences. Mol.Cell Biol. 9: 4248-4258.
- Chi, T.H. and Crabtree, G.R. 2000. Perspectives: signal transduction. Inositol 14. phosphates in the nucleus [comment]. Science 287: 1937-1939.
- Clark, A.J., Archibald, A.L., McClenaghan, M., Simons, J.P., Wallace, R., and 15. Whitelaw, C.B. 1993. Enhancing the efficiency of transgene expression. Philos.Trans.R.Soc.Lond B Biol.Sci. 339: 225-232.
- Clements, S., Mehansho, H., and Carlson, D.M. 1985. Novel multigene families 16. encoding highly repetitive peptide sequences. Sequence analyses of rat and mouse proline-rich protein cDNAs. J Biol.Chem. 260: 13471-13477.
- Corring, T. 1980. Endogenous secretions in the pig. In Current concepts of digestion 17. and absorption in pigs. Edited by A.G.Low and I.G.Partridge. National Institute for Research in Dairying, Reading pp. 136-150.
- Cosgrove, D.J. 1980. Inositol phosphates. Their chemistry, biochemistry and 18. physiology. Elsevier, Amsterdam.
- Crameri, A., Dawes, G., Rodriguez, E.Jr., Silver, S. and Stemmer, W.P.C. 1997. 19. Molecular evolution of an arsenate detoxification pathway by DNA shuffling. Nature Biotechnol. 15: 436-438.
- Dassa, J., Marck, C., and Boquet, P.L. 1990. The complete nucleotide sequence of the 20: Escherichia coli gene APPA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase. J.Bacteriol. 172: 5497-5500.

- Denny, H.R. and Messervy, A. 1972. Surgical techniques for the extirpation of the submandibular salivary glands and the collection of salivary secretions in the pig. Vet.Rec. 90: 650-654.
- 22. Dobrinsky, J.R., Johnson, L.A., and Rath, D. 1996. Development of a culture medium (BECM-3) for porcine embryos: Effect of bovine serum albumin and fetal bovine serum on embryo development. Biol. Reprod. 55: 1069-1074.
- Dvorakova, J. 1998. Phytase: sources, preparation and exploitation. Folia Microbiol. (Praha) 43: 323-338.
- 24. Engelen, A.J., van der Heeft, F.C., Randsdorp, P.H., and Smit, E.L. 1994. Simple and rapid determination of phytase activity. J AOAC Int. 77: 760-764.
- Frandson, R.D. and Spurgeon, T.L. 1992. Anatomy and physiology of farm animals.
 Lea & Febiger, Philadelphia.
- Galfre, G. and Milstein, C. 1981. Preparation of monoclonal antibodies: strategies and procedures. Methods Enzymol. 73: 3-46.
- 27. Geyer, P.K. 1997. The role of insulator elements in defining domains of gene expression. Curr. Opin. Genet. Dev. 7: 242-248.
- 28. Gish, W. and States, D.J. 1993. Identification of protein coding regions by database similarity search. Nature Genetics 3: 266-272.
- Golovan, S., Wang, G., Zhang, J., and Forsberg, C.W. 2000. Characterization and overproduction of the Escherichia coli APPA encoded bifunctional enzyme that exhibits both phytase and acid phosphatase activities [In Process Citation].
 Can. J. Microbiol. 46: 59-71.
- Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, J.A., and Ruddle, F.H. 1980.
 Genetic transformation of mouse embryos by microinjection of purified DNA.
 Proc.Natl.Acad.Sci.USA 77: 7380-7384.

- Gorman, C., Padmanabhan, R., and Howard, B.H. 1983. High efficiency DNAmediated transformation of primate cells. Science 221: 551-553.
- Greiner, R. and Jany, K.-D. Characterization of a phytase from Escherichia coli.
 Herbsttagung der Gesellschaft fur Biologische Chemie 372. 1991.
 Ref Type: Abstract
- Greiner, R., Konietzny, U., and Jany, K.D. 1993. Purification and characterization of two phytases from Escherichia coli. Arch Biochem. Biophys. 303: 107-113.
- 34. Hall, J., Hazlewood, G.P., Surani, M.A., Hirst, B.H., and Gilbert, H.J. 1990. Eukaryotic and prokaryotic signal peptides direct secretion of a bacterial endoglucanase by mamalian cells. J.Biol.Chem. 275: 19996-19999.
- 35. Harland, B.F. and Morris, E.R. 1995. Phytate: A good or a bad food component? Nutr. Res. 15: 733-754.
- Harlow, E., and Lane, D. 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, New York.
- Harayama, S. 1998. Artificial evolution by DNA shuffling. Trends Biotechnol. 16:
 76-81.
- 38. Heinoen, J.K. and Lahti, R.J. 1981. A new and convenient colorimetric determination of inorganic orthophosphate and its application to the assay of inorganic pyrophosphate. Anal. Biochem. 133: 313-317.
- Heneine, W. and Switzer, W.M. 1996. Highly sensitive and specific polymerase chain reaction assays for detection of baboon and pig cells following xenotransplantation in humans. Transplantation 62: 1360-1362.
- 40. Higuchi, R. 1989. Simple and rapid preparation of samples for PCR. In PCR technology: principles and applications for DNA amplification. Edited by H.A. Erlich. Macmillan Publishers; Stockton Press, New York, NY pp. 31-38.

- 41. Hu,Y., Nakagawa,Y., Purushotham,K.R., and Humphreys-Beher,M.G. 1992. Functional changes in salivary glands of autoimmune disease-prone NOD mice. Am.J.Physiol. 263: E607-E614.
- 42. Huang, X. 1996. An improved sequence assembly program. Genomics 33: 21-31.
- 43. Huang, X. 1999. A contig assembly program based on sensitive detection of fragment overlaps. Genomics 14: 18-25.
- 44. Huff, W.E., Moore, P.A., Jr., Waldroup, P.W., Waldroup, A.L., Balog, J.M., Huff, G.R., Rath, N.C., Daniel, T.C., and Raboy, V. 1998. Effect of dietary phytase and high available phosphorus corn on broiler chicken performance. Poult Sci 77: 1899-1904.
- 45. Iqbal, T.H., Lewis, K.O., and Cooper, B.T. 1994. Phytase activity in the human and rat small intestine. Gut 35: 1233-1236.
- 46. Jia,Z., Golovan,S., Ye,Q., and Forsberg,C.W. 1998. Purification, crystallization and preliminary X-ray analysis of the Escherichia coli phytase. Acta Crystallogr.D.Biol.Crystallogr. 54 (Pt 4): 647-649.
- 47. Kazazian, H.H.J. 1998. Mobile elemants and disease. Curr. Opin. Genet. Dev. 8: 343-350.
- 48. Kim, H.S. and Maeda, N. 1986. Structures of two HaeIII-type genes in the human salivary proline-rich protein multigene family. J Biol. Chem. 261: 6712-6718.
- 49. Kohler, G. and Milstein, C. 1976. Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. Eur. J. Immunol. 7: 511-519.
- Kozak, M. 1987. At least six nucleotides preceding the AUG initiator codon enhance traslation in mammalian cells. Proc. Natl. Acad. Sci. USA 196: 947-950.
- 51. Kuchner, O. and Arnold, F.H. 1997. Directed evolution of enzyme catalysts. Trends Biotechnol. 15: 523-530.
- 52. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-785.

- 53. Larsen, H.J., Brodersen, C.H., and Hjorth, J.P. 1994. High-level salivary gland expression in transgenic mice. Transgenic Res. 3: 311-316.
- 54. Laursen, J. and Hjorth, J.P. 1997. A cassette for high-level expression in the mouse salivary glands. Gene 198: 367-372.
- 55. Levitt, N., Briggs, D., Gil, A., and Proudfoot, N.J. 1989. Definition of an efficient synthetic poly(A) site. Genes Dev. 3: 1019-1025.
- 56. Low, A.G. 1989. Research into the digestive physiology of pigs. In Nutrition and digestive physiology in monogastric farm animals. Edited by E.J.van Weerden and J.Huisman. Pudoc, Wageningen pp. 1-15.
- 57. Madsen, H.O. and Hjorth, J.P. 1985. Molecular cloning of mouse PSP mRNA. Nucleic Acids Res. 13: 1-13.
- 58. Mallin, M.A. 2000. Impacts of industrial animal production on rivers and estuaries.

 American Scientist 88: 26-37.
- Mehansho, H., Ann, D.K., Butler, L.G., Rogler, J., and Carlson, D.M. 1987. Induction of proline-rich proteins in hamster salivary glands by isoproterenol treatment and an unusual growth inhibition by tannins. J.Biol.Chem. 262: 12344-12350.
- 60. Menniti, F.S., Oliver, K.G., Putney, J.W., Jr., and Shears, S.B. 1993. Inositol phosphates and cell signaling: new views of InsP5 and InsP6. Trends Biochem. Sci 18: 53-56.
- Mikkelsen, T.R., Brandt, J., Larsen, H.J., Larsen, B.B., Poulsen, K., Ingerslev, J., Din, N., and Hjorth, J.P. 1992. Tissue-specific expression in the salivary glands of transgenic mice. Nucleic Acids Res. 20: 2249-2255.
- 62. Miner, J.R. 1999. Alternatives to minimize the environmental impact of large swine production units. J.Anim Sci. 77: 440-444.
- 63. Nakai, K. and Kanehisa, M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14: 897-911.

- 64. Nesterenko, M.V., Tilley, M., and Upton, S.J. 1994. A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in polyacrylamide gels. J Biochem. Biophys. Methods 28: 239-242.
- O'Shannessy, D.J., Voorstad, P.J., and Quarles, R.H. 1987. Quantitation of glycoproteins on electroblots using the biotin-streptavidin complex. Anal. Biochem. 163: 204-209.
- 66. Owerbach, D. and Hjorth, J.P. 1980. Inheritance of a parotid secretory protein in mice and its use in determining salivary amylase quantitative variants. Genetics 95: 129-141.
- 67. Pen,J., Verwoerd,T.C., Vanparidon,P.A., Beudeker,R.F., Vandenelzen,P.J.M., Geerse,K., Vanderklis,J.D., Versteegh,H.A.J., Vanooyen,A.J.J., and Hoekema,A. 1993. Phytase-containing transgenic seeds as a novel feed additive for improved phosphorus utilization. Biotechnology 11: 811-814.
- 68. Petitclerc, D., Attal, J., Theron, M.C., Bearzotti, M., Bolifraud, P., Kann, G., Stinnakre, M.G., Pointu, H., Puissant, C., and Houdebine, L.M. 1995. The effect of various introns and transcription terminators on the efficiency of expression vectors in various cultured cell lines and in the mammary gland of transgenic mice. J Biotechnol. 40: 169-178.
- 69. Pinkert, C.A., Dyer, T.J., Kooyman, D.L., and Kiehm, D.J. 1990. Characterization of transgenic livestock production. Domest. Anim Endocrinol. 7: 1-18.
- Quissell, D.O., Barzen, K.A., Redman, R.S., Camden, J.M., and Turner, J.T. 1998.
 Development and characterization of SV40 immortalized rat parotid acinar cell lines.
 In Vitro Cell Dev. Biol. Animal. 34: 58-67.
- 71. Reddy, N.R., Sathe, S.K., and Salunkhe, D.K. 1982. Phytates in legumes and cereals. Adv. Food Res. 28: 1-92.
- 72. Rozhkov, Y.I. and Galimov, I.R. 1990. Salivary gland polymorphism in pigs and cattle detected by affinity electrophoresis. Anim. Genet 21: 277-283.

- Samuelson, L.C. 1996. Transgenic approaches to salivary gland research.
 Annu. Rev. Physiol. 58: 209-229.
- 74. Shaw, P. and Schibler, U. 1986. Structure and expression of the parotid secretory protein gene of mouse. J Mol. Biol. 192: 567-576.
- 75. Simons, P.C., Versteegh, H.A., Jongbloed, A.W., Kemme, P.A., Slump, P., Bos, K.D., Wolters, M.G., Beudeker, R.F., and Verschoor, G.J. 1990. Improvement of phosphorus availability by microbial phytase in broilers and pigs. Br J Nutr. 64: 525-540.
- Smit,A.F.A. 1996. The origin of interspersed repeats in the human genome.
 Curr.Opin.Genet.Dev. 6: 743-748.
- 77. Stahl, C.H., Han, Y.M., Roneker, K.R., House, W.A., and Lei, X.G. 1999. Phytase improves iron bioavailability for hemoglobin synthesis in young pigs. J Anim Sci 77: 2135-2142.
- 78. Svendsen, P., Laursen, J., Krogh-Pedersen, H., and Hjorth, J.P. 1998a. Novel salivary gland specific binding elements located in the PSP proximal enhancer core. Nucleic Acids Res. 26: 2761-2770.
- 79. Taboit-Dameron, F., Malassagne, B., Viglietta, C., Puissant, C., Leroux-Coyau, M., Chercau, C., Attal, J., Weill, B., and Houdebine, L.M. 1999. Association of the 5'HS4 sequence of the chicken beta-globin locus control region with human EF1 alpha gene promoter induces ubiquitous and high expression of human CD55 and CD59 cDNAs in transgenic rabbits. Transgenic Res. 8: 223-235.
- 80. Terada, T. and Nakanuma, Y. 1993. An immunohistochemical survey of amylase isoenzymes in cholangiocarcinoma and hepatocellular carcinoma. Arch Pathol. Lab Med. 117: 160-162.
- 81. Towbin, H., Staehelin, T., and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc.Natl.Acad.Sci.USA 76(9), 4350-4354. 1979.
 Ref Type: Generic

- Tryon, A.F. and Bibby, B.G. 1966. Preliminary studies on pig saliva. Arch. Oral Biol. 82. 11: 527-531.
- Tu, Z.J., Lazowski, K.W., Ehlenfeldt, R.G., Wu, G., Lin, H.H., Kousvelari, E., and 83. Ann, D.K. 1993. Isoproterenol/tannin-dependent R15 expression in transgenic mice is mediated by an upstream parotid control region. Gene Expr. 3: 289-305.
- Verwoerd, T.C., van Paridon, P.A., van Ooyen, A.J., van Lent, J.W., Hoekema, A., and 84. Pen, J. 1995. Stable accumulation of Aspergillus niger phytase in transgenic tobacco leaves. Plant Physiol 109: 1199-1205.
- Wall, R.J., Hyman, P., Kerr, D., Pintado, B., and Wells, K. 1997. Transgenic animal 85. technology. J.Androl 18: 236-239.
- Wall, R.J., Pursel, V.G., Hammer, R.E., and Prinster, R.L. (Editors). 1985. Development 86. of porcine ova that were centrifuged to permit visualization of pronuclei and nuclei. Biol.Reprod. 32.
- Wessel, D. and Flugge, U.I. 1984. A method for the quantitative recovery of protein 87. in dilute solution in the presence of detergents and lipids. Anal. Biochem. 138: 141-143.
- Whitelaw, C.B., Archibald, A.L., Harris, S., McClenaghan, M., Simons, J.P., and 88. Clark, A.J. 1991. Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice. Transgenic Res. 1: 3-13.
- Williamson, C.M., A.J.Bramley, and A.J.Lax 1994. Expression of the lysostaphin gene 89. of Staphylococcus simulans in a eukaryotic system. Appl. Environ. Microbiol. 60: 771-*7*76.
- Wodzinski, R.J. and Ullah, A.H. 1996. Phytase. Adv. Appl. Microbiol. 42: 263-302. 90.
- Xiong, Y. and Eickbush, T.H. 1990. Origin and evolution of retroelements based upon 91. their reverse transcriptase sequences. EMBO J. 9: 3353-3362.

25

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

- A transgenic non-human animal that carries in the genome of its somatic and/or germ
 cells a nucleic acid sequence including a heterologous transgene construct, said construct including a trangene encoding a protein, said transgene being operably linked to a first regulatory sequence for salivary gland specific expression of said protein.
- The animal of claim 1 wherein said first regulatory sequence comprises a saliva
 protein promoter/enhancer sequence, whereby said animal expresses said protein in its saliva.
 - 3. The animal of claim 1 wherein said animal is a mammal.
- 4. The animal of claim 3 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats and dogs, and in addition, fish and poultry.
 - 5. The animal of claim 1 wherein said saliva protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.

6. The animal of claim 5 wherein said promoter/enhancer is a parotid secretory protein (PSP) promoter/enhancer.

- 7. The animal of claim 6 wherein said parotid secretory protein (PSP) promoter/enhancer is derived from a mouse.
- 8. The animal of claim 5 wherein said promoter/enhancer is a proline-rich protein (PRP) promoter/enhancer.
- 30 9. The animal of claim 8 wherein said proline-rich protein (PRP) promoter/enhancer is derived from a rat.

- 10. The animal of claim 1 wherein said transgene is further operably linked to one or more second regulatory sequences including enhancers, transcription regulatory sequences, termination sequences, and polyadenylation sites.
- 5 11. The animal of claim 1 wherein said transgene comprises a gene encoding a protein having phytase activity.
 - 12. The animal of claim 1 wherein said transgene encodes a phytase or a homologue thereof.

13. The animal of claim 1 wherein said animal is a pig, said transgene comprising a gene encoding a protein having phytase activity and wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer or a proline-rich protein (PRP) promoter/enhancer.

15

- 14. The animal of claim 1 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.
- 15. A transgenic non-human animal that carries in the genome of its somatic and/or germ
 20 cells a nucleic acid sequence including a heterologous transgene construct, said construct including a trangene encoding phytase or a homologue thereof.
 - 16. The animal of claim 15 wherein said transgene is operably linked to a first regulatory sequence for salivary gland specific expression of said phytase.

- 17. The animal of claim 16 wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.
- 30 18. The animal of claim 17 wherein said animal is a mammal.

- 19. The animal of claim 18 wherein said phytase or a homologue thereof is expressed in saliva or in the gastrointestinal tract of said animal.
- The animal of claim 15 wherein said transgene construct comprises a nucleic acid
 sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.
 - 21. A method of expressing a protein, the method comprising the steps of:
 - a) introducing a transgene construct into a non-human animal embryo such that a non-human transgenic animal that develops from said embryo has a genome that comprises said transgene construct, wherein said transgene construct comprises:
 - i) a transgene encoding said protein, and
 - ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein,
 - b) transferring said embryo to a foster female; and,
- c) developing said embryo into said transgenic animal
 wherein said transgene is produced in the gastrointestinal tract of said animal.
 - 22. The method of claim 21 wherein said regulatory sequence provides for salivary gland or pancreatic gland specific expression of said protein.
 - 23. The method of claim 21 wherein said regulatory sequence provides for salivary gland specific expression of said protein.
- The method of claim 23 wherein said salivary gland is a parotid gland, submaxillarygland, or a submandibular gland.
 - 25. The method of claim 23 wherein said transgene is expressed in the saliva of said animal.
- 30 26. The method of claim 21 wherein said transgene is heterologous.

60

- 27. The method of claim 21 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.
- 28. The method of claim 21 wherein said protein is a glycoprotein.
- 29. A transgenic animal adapted for expressing a protein according to the method of claim 21, or a progeny thereof.
- 30. The method of claim 21 wherein said protein is a phytase or a homologue thereof.
- 31. The method of claim 21 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:7.
- 32. A process for producing a protein comprising the steps of:
- a) obtaining saliva containing said protein from a non-human transgenic animal, said animal containing within its genome a transgene construct, wherein said transgene construct comprises:
 - i) a transgene encoding said protein, and
 - ii) at least one regulatory sequence for salivary gland specific expression of said protein, and

extracting said protein from said saliva.

5

10

- 33. The process of claim 32 wherein said transgene is heterologous.
- 25 34. The process of claim 32 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.
 - 35. The process of claim 32 wherein said protein is a glycoprotein.
- 36. The process of claim 32 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

- 37. The process of claim 32 wherein said protein is a phytase or a homologue thereof.
- 38. The process of claim 32 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

10

15

- 39. A method for expressing a phytase or a homologue thereof in a non-human animal, said method comprising:
 - a) constructing a nucleic acid sequence including a transgene construct comprising:
 - i) a transgene encoding said phytase or a homologue thereof, and
 - ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein, and
- b) transfecting the animal with said nucleic acid sequence; whereby said animal carries within the genome of its somatic and/or germ cells said transgene construct and wherein said animal expresses said phytase or a homologue thereof in its gastrointestinal tract.
- 40. The method of claim 39 wherein said transgene construct results in salivary gland or pancreatic gland specific expression of said phytase or a homologue thereof.
- 20 41. The method of claim 40 wherein said regulatory sequence provides for salivary gland specific expression of said phytase or a homologue thereof.
 - 42. The method of claim 41 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

- 43. The method of claim 41 wherein said phytase or a homologue thereof is expressed in the saliva of said mammal.
- 44. The method of claim 41 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

20

25

- 45. The method of claim 39 wherein said nucleic acid sequence is introduced into said animal in the form of a transgene construct.
- 46. The method of claim 45 wherein said transgene construct is a nucleic acid molecule.
- 47. The method of claim 46 wherein said plasmid comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:6.
- 48. The method of claim 39 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats, dogs, fish and poultry.
 - 49. The method of claim 48 wherein said animal comprises a mouse or a pig.
- 50. A nucleic acid molecule comprising a nucleic acid sequence including a gene
 encoding a protein, said gene being operably linked to at least one regulatory sequence for gastrointestinal tract specific expression of said protein.
 - 51. The molecule of claim 50 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence, whereby expression of said protein is salivary gland specific.
 - 52. The molecule of claim 51 wherein said salivary protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer, a salivary amylase promoter/enhancer, or a SV40 promoter/enhancer.
 - 53. The molecule of claim 51 wherein said protein comprises a phytase or a homologue thereof.
 - 54. The molecule of claim 53 wherein said molecule is a transgene construct.
 - 55. The molecule of claim 54 wherein said molecule is a nucleic acid molecule.

63

- 56. The molecule of claim 55 wherein said molecule comprises a nucleic acid sequence according to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6.
- 57. The molecule of claim 53 wherein said molecule includes a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.
 - 58. An antibody specific to a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.
- 10 59. The antibody of claim 58 wherein said antibody is monoclonal.
 - 60. The antibody of claim 58 wherein said antibody is polyclonal.
 - 61. A hybridoma secreting the antibody of claim 59.
- 62. A host cell transfected with molecule of claim 50.

15

- 63. A host cell transfected with the molecule of claim 56.
- 20 64. A host cell transfected with the molecule of claim 57.
 - 65. The host cell of claim 63 wherein said cell is an bacterial cell.
 - 66. The host cell of claim 64 wherein said cell is an animal cell.
 - 67. A diagnostic kit for immunologically detecting a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7, the kit including an antibody specific to said protein.
- 30 68. The kit of claim 67 wherein said antibody is monoclonal.
 - 69. The kit of claim 68 wherein said antibody is polyclonal.

WO 00/64247

ġ.

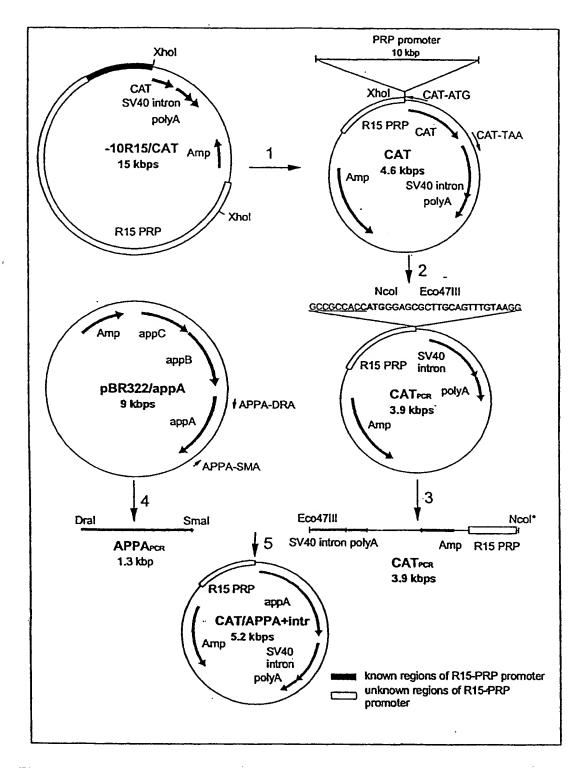


Figure 1

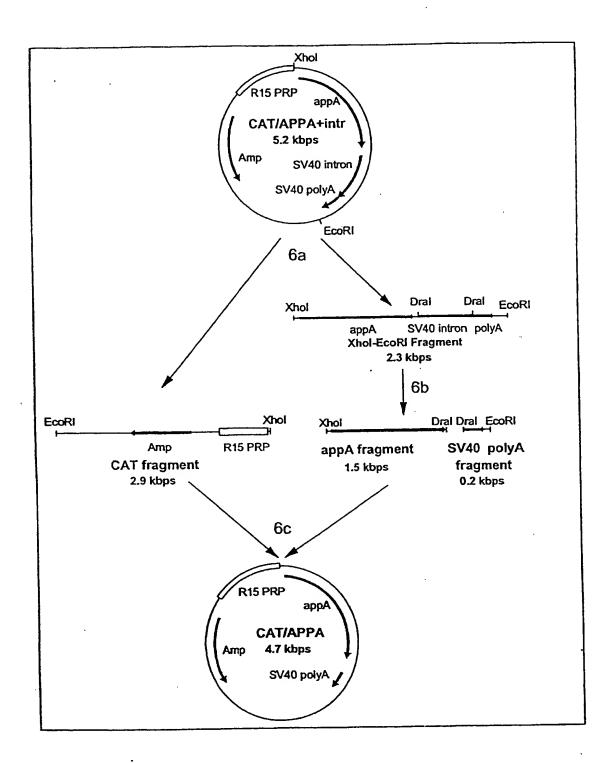


Figure 1 (continued)

2/58
RECTIFIED SHEET (RULE 91)

WO 00/64247

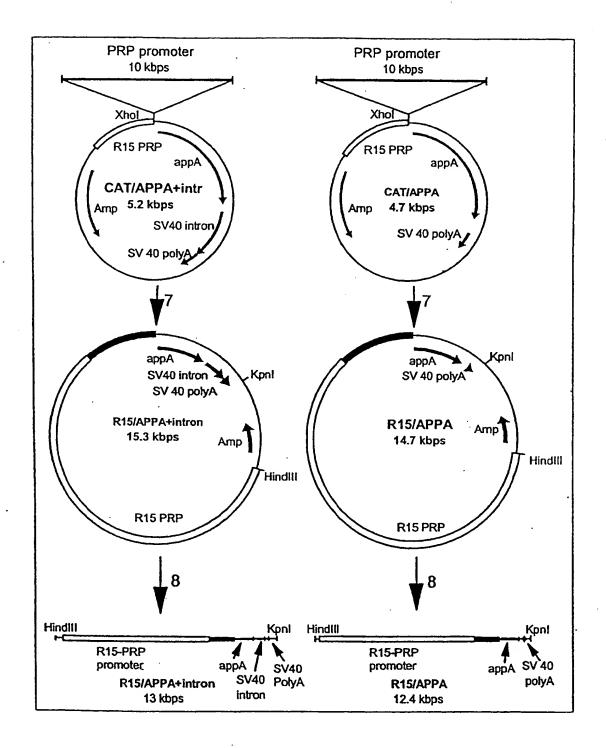


Figure 1 (continued)

3/58
RECTIFIED SHEET (RULE 91)

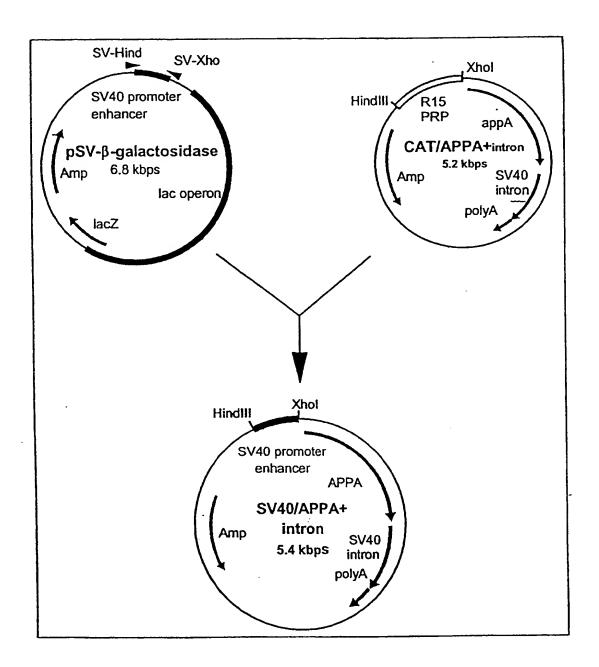


Figure 2

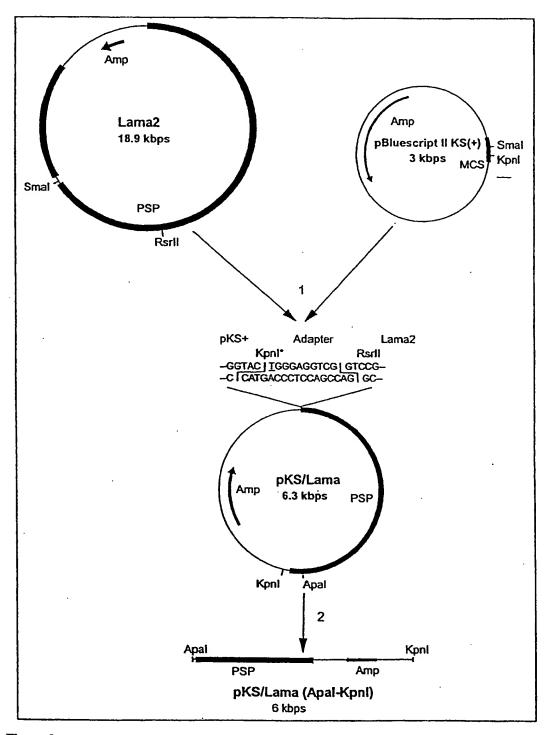


Figure 3

5/58
RECTIFIED SHEET (RULE 91)

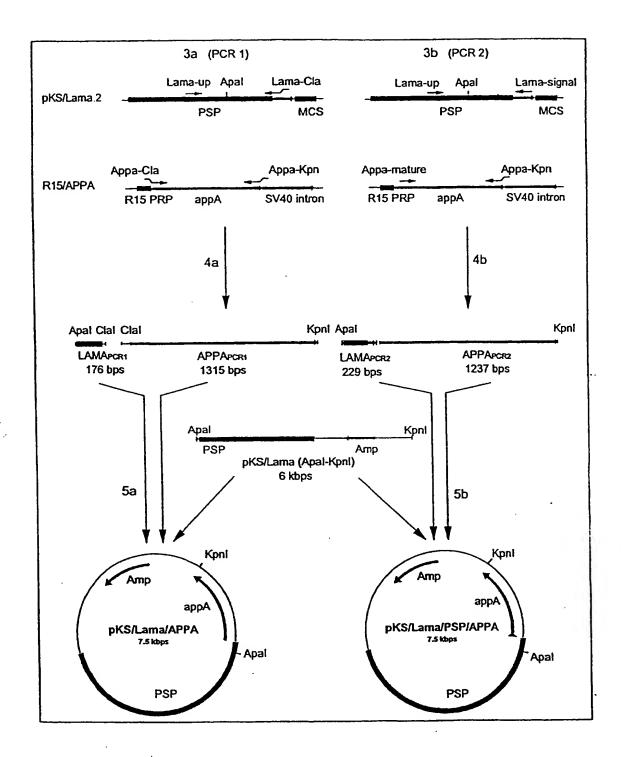


Figure 3 (continued)

6/58
RECTIFIED SHEET (RULE 91)

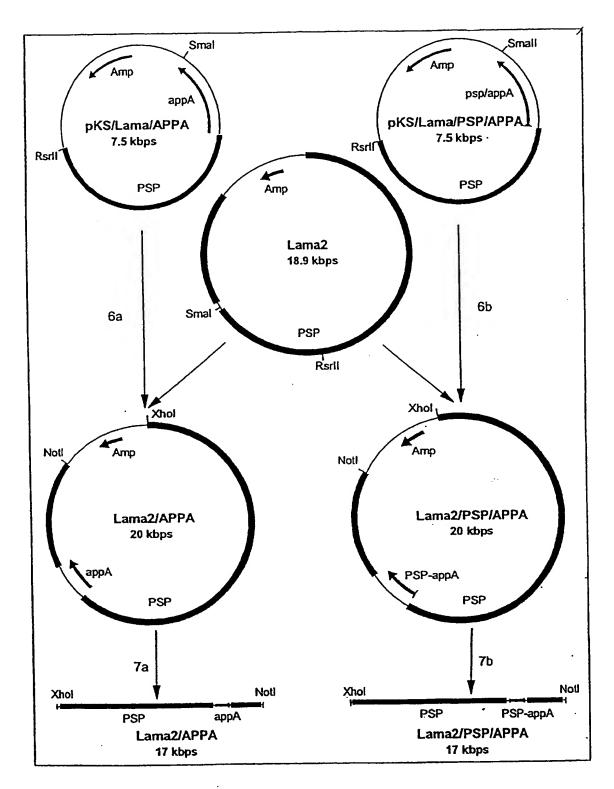


Figure 3 (continued)

7/58
RECTIFIED SHEET (RULE 91)

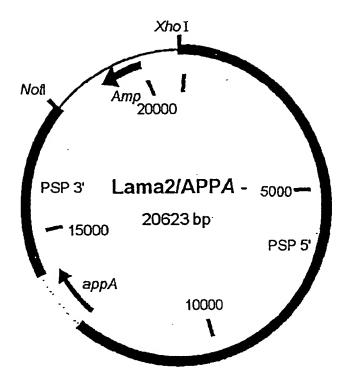


Figure 4. Schematic diagram of the Lama2/APPA construct.

Figure 5. The nucleic acid sequence of the Lama2/APPA plasmid (SEQ ID NO: 1)

DNA CIRCULAR SYN 17-JAN-2000 20623 bp LOCUS Lama-appA Lama 2/APPA transgenic construct DEFINITION ACCESSION Lama 2-appA, parotid secretory protein; acid glucose-1-phosphatase; appA KEYWORDS periplasmic phosphoanhydride phosphohydrolase; artificial sequence: cloning vector REFERENCE 1 (bases 1 to 20623) Golovan, S., Forsberg, C.W., Phillips, J. AUTHORS JOURNAL Unpublished. FEATURES DEFINITION M. musculus Psp gene for parotid secretory protein. ACCESSION X68699 X68699.1 GI:53809 SOURCE house mouse. ORGANISM Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. REFERENCE 1 (bases 3777 to 5332;) AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P. Novel salivary gland specific binding elements located in the PSP TITLE proximal enhancer core Nucleic Acids Res. 26 (11), 2761-2770 (1998) JOURNAL 98256451 MEDLINE 2 (bases 7147 to 12653; 13952 to 17731) REFERENCE AUTHORS Mikkelsen, T.R. TITLE Direct Submission Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular JOURNAL Biology, University of Aarhus, CF Mollers Alle 130, 8000 Aarhus, DENMARK REFERENCE 3 (bases 7147 to 12653; 13952 to 17731) AUTHORS Laursen J, Hjorth JP TITLE A cassette for high-level expression in the mouse salivary glands. Gene 1997 Oct 1;198(1-2):367-72 JOURNAL MEDLINE 9370303 Location/Qualifiers **FEATURES** 1.to 12653; 13952 to 17731 source /organism="Mus musculus" /strain="C3H/As" /db_xref="taxon:10090" /chromosome="2" /map="Estimate: 69 cM from centromere" /clone="Lambda YP1, Lambda YP3, Lambda YP7" /clone lib="Lambda-PHAGE (Lambda L47.1)" /germline /note="Allele: b" 3777-5332 misc_feature /gene="PSP" /function="salivary gland specific positive acting regulatory region" 7147..8724 enhancer /evidence=experimental exon 11778..11824 /gene="Psp" /note="exon a" /number=1 /evidence=experimental 12626.. 14190 exon /gene="Psp" /note="exon b fused with exons h and i" misc_feature 12644-12652

Figure 5 (continued): /function=" consensus sequence for initiation in higher eukaryotes misc_feature 13952-13965 /function=" M13mp18 polylinker" DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene, M58708 L03370 L03371 L03372 L03373 L03374 L03375 ACCESSION M58708.1 GI:145283 VERSION Escherichia coli DNA. SOURCE Escherichia coli ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia. REFERENCE 1 (bases 12653..13951) Dassa, J., Marck, C. and Boquet, P.L. AUTHORS The complete nucleotide sequence of the Escherichia coli gene appA TITLE reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase J. Bacteriol. 172 (9), 5497-5500 (1990) JOURNAL MEDLINE 90368616 Location/Qualifiers **FEATURES** 12653..13951 Source /organism="Escherichia coli" /db_xref="taxon:562" 12653..12718 sig_peptide /gene="appA" CDS12653 13951 /gene="appA" /standard_name="acid phosphatase/phytase" /transl table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285" translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP/ TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS GOVALIADVDERTRKTGEAFAAGLAPDCALTVHTQADTSSPDPLFNPLKTGVCQLDNA nvtdailsraggsiadftghrqtafrelervlnfpqsnlclkrekqdescsltqalps ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF YLLQRTPEVARSRATPLLIDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL* 12719 13948 mat_peptide /gene="appA" /product="periplasmic phosphoanhydride phosphohydrolase" replace(12659.. 12661, "gcg changed to gcc") mutation /gene="appA" /standard_name="A3 mutant" /note="created by site directed mutagenesis" /citation=[3] /phenotype="silent mutation" replace(13934..13936, " ccg changed to ccc") mutation /gene="appA" /standard name= " P428 mutant" /note="created by site directed mutagenesis" /citation=[3] /phenotype=" silent mutation " replace (13937..13939, * gcg changed to gct*) mutation

/note="created by site directed mutagenesis"

10/58

/standard_name=" A429 mutant"

/gene="appA"

Figure 5 (continued):

```
/citation=[3]
                   /phenotype=" silent mutation "
      DEFINITION
                   pBluescript II KS(+) vector DNA.
      ACCESSION
                   X52327
      VERSION
                   X52327.1 GI:58061
      KEYWORDS
                   artificial sequence; cloning vector; expression vector; vector.
      SOURCE
                   synthetic construct.
      ORGANISM
                   synthetic construct
                   artificial sequence.
      REFERENCE
                         (bases 17732 to 20623)
      AUTHORS
                   Thomas, E.A.
      TITLE
                   Direct Submission
      JOURNAL
                   Submitted (20-FEB-1990) Thomas E.A., Stratagene Cloning
                   Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
      REFERENCE
                         (bases 17732 to 20623)
                   Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D.
      AUTHORS
      TITLE
                   Lambda ZAP: a bacteriophage lambda expression vector with in
                   vivo excision properties
      JOURNAL.
                   Nucleic Acids Res. 16 (15), 7583-7600 (1988)
      MEDLINE
                   88319944
      REFERENCE
                         (bases 17732 to 20623)
                   3
      AUTHORS
                   Alting-Mees, M.A. and Short, J.M.
      TITLE
                   pBluescript II: gene mapping vectors
                   Nucleic Acids Res. 17 (22), 9494 (1989)
      JOURNAY.
      MEDLINE
                   90067967
      FEATURES
                         Location/Qualifiers
      Source
                         17732 to 20623
                         /organism="synthetic construct"
                         /db xref="taxon:32630"
      CDS
                         complement (18967..19827)
                   /gene="Amp"
                   /product="b-lactamase"
BASE COUNT
              5449 a
                       4847 C
                               4902 g 5424 t
ORIGIN
       1 TCGAGAGTAT CTTTGTCAGC TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAG
      61 ATCTARACTA ATTARTTART CCCTCACCCG CARATCTTC AGTCACTARG TTAGCACGAT
     121 TGTTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
     181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
     241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
     301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
     361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
     421 AGGGATTGAG AGACCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACTTG
     481 ACAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGACCCG
     541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCAGTATT
     601 ATAAAAGGAC AGTATTACAG ATTTTGTTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
     661 TITAAGTAGG GTAAAGTACT CITTAAAAAT GGGTCCTAGA TAITITTTCC TITAACTCAA
     721 GICTCITACT GITTAAATGA TITTTATTIT GITTAATATG GAGGAAAAAG AAGCGTAAAT
     781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
     841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
     961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
    1021 GCGGGGGTG GTGGCATACA CTTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
    1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
    1201 ACCAAACCAA ACCAAACCAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
    1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
    1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACCAGCCTG AGTGTCCATG ATAAATGAAT
    1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTTAT
    1441 TCATTTTCT TTATGAGGIG TCCATTCAGG AGTCACATGG TAGITCTATT TTCAGTCTTC
    1501 TGAAGATACT ACACTGGTCC CCACAGTTTA CACTTTTATC AGCAGTGAAT AAGGGTTCCT
    1561 CTATCCTTAC CATCATTTGT TGTAATTTTT CTTGATGACC CTCTTTCTGA CAGGGATAGG
    1621 ATGTAATATC AGTGTGAGGA AGTACAACTT GTTTTCTAAG TATTTATTGG CCCCTTGCAT
    1681 TTCTTCTTTT GAAAACTGTC GGTTCCTGAC ATCTGCTCAG GTATTCATTG GATGTTGTTT
                                     11/58
```

Figure 5 (continued):

1741 CTTTGGTGTT TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC 1801 ATTCTGTAGG CTGCCTCCTC ACCCTGGCAA TTGTTGTCCT TGTTTTGCAG AAACTTTTGA 1861 CTTCATGGAA TCTCATTTGT CAGTTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG 1921 GTTTTTACAG AGCCCTGGTC TATGCCTTTA TCCTCCTCTG GCAGCTTCGG AGTTTCATTT 1981 CTTACATTTA GATCTTTGAT CCACTTTGAA CAAGTTTTGG AGCAGGGTGA GAGATACGAA 2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTTACATAG CATCGTTGGT TGAAGAGGTT 2101 TTATTTTATT TTTAAATAAT GTGTCATAAA AAACGAGGTG GTTGTAGCAG TGTGGATTTG 2161 TITCTITGTC CITTGATCTA CAGGICITGT TITGTGTCAG TCTCATGATG TITTATTGCT 2221 ATGCTCTGT CATACAGTCT GAGGTCAGGT ATTGTGATAT ACCTTCAGTA TTGCTCCCTC 2281 AGACTCAGGT TTGCTTTGGC CAGGAGTCAT CTTACTCAGT GCTCTTAGAG CTCCCCCAGC 2341 ATGTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAACT 2401 GTCTTGAACT ACTTCTGGGG AGGTGAAACG TGGAGACACT AAACTGTGTT TACCCTGTAC 2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGGT TCCCAAGCTC CCAGGTGTCC AGTCACTGCT 2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC 2641 TGGGGCTGGG ATCTCCCAAA CACCTGGATA TCCAGATGCC ACTGGGTCAG GGGGAGTTGG 2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC 2761 TGGCTTTACT AATTTGCGAA AGTCCTTAGC TTAGCAGCAG TTGTCTGGGA GCACAGAGGG 2821 GCCTTCTGTA AGAGGCTCAG GCAGTGCCGC TCTGTAGGCG AAGGTCTTCT CCATGTTCCC 2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCCAA ACTGGTTTAT TGATTGTTCA 2941 TTGTGGAAAA TGGGTGCACA CCACCTTCTC AGGGTGGACC AGAGATCAAA TACCTTTTGC 3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACAT 3061 CATTAGCATG GAGAACTCTG TTCTGGGCTA CATGACCACA GGCCACATTT CCACAAGCCA 3121 CATGTGGGAA GTGTGGCACA TGTTCTAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC 3181 TACCATCCCA GGTGGGTGCC TGGGTGCCAG GGACCCTGAA CCCGCTCAAC CTTACCAAGT 3241 TTCCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTGTCTTT 3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTTCTCCACC TGTACCTCAT TGGTTTGGGT 3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAA GCATCGTGGC 3421 AGTCCAGGAA CCACACCCTG AGGTTCCTGC ACTGAAGGGA CTCCCTAAGT CTCTGGAGTC 3481 TCTCCCCTTC ACAGAGCTGC CAAAGTCTAG GTTCTTTTGA GGATAACAGA GCCATGCTTG 3541 GTAAGCAGAC AACAGCATTT GTTTACTCAA CCTTCTTTTG TCAGCTCCCT CTTCATAAAC 3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAACT GTGCAAGGAA 3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG 3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT 3781 CCTGGGAACC CCTCTACCAT GACACACATT CTTCCTGCAG GTCACACTTA ATAGGCCATT 3841 TCTTATTTGG ATCTATCATG GTGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGCGAACA 3901 GAAAGTTATA TAAAAACAAG TCCCCCCCC TTGTCACTGC TGCTAAGAAT GTAGCAGAAA 3961 TIGICTCAAG IGICTCTCTA ATCAGAAACA ATAAAGGTCT CCTIGGATTC AAGCCCTCCA 4021 GTTTCCTCCT TCCTTGCTGA GCCTTGGACA CCCATACAAA CCTCCTGGAT GCTACAGCTC 4081 TGGGCAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAAATA CTTGTTTGGG 4141 GGTACACCCA CTCCTCTGCC TGTGTGGTTC CTGCAGTCAG TCCTGCAGAC AGGCCCTCAG 4201 TGGGTCTTCC ATGGGCAACA CGCAGAGGGA GGCAATGGAT GGGAATACCC ACACCCTGGT 4261 TAGITTACCC CGGCCATGCT CTCTGCTCTT CATCCCTCCT CTGCCCTCTG CCACGGCTTT 4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCCTCACCC TAGCTATGAT 4381 GTITACTITA GAGTGACCIT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCACGGAG 4441 ATGCTAGGGA AGAAACGTCT TCTAACTACT GAGGTTACTA AGTTCCTGGT GGTTGTCTCT 4501 GCCTTTCCCT TGTTAAAGTC ACCTTGAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA 4561 GAGTAAATAT GGTCCTGAAG ATTTCCTTTG AGTGCCCAGA ATCCATGACA TTTCAAGAGC 4621 CCTCTTGTA CCTTAAGTCA TITGGGGTTG TATCTTCTGC TIGATGTATG TGTGTGTGT 4681 TATCARAGAG TGAGATGGTT ACATARGAGG TGCTCTARAG GACAGAGAG ATTTGCARTT 4741 GTGCATGTG ACATCCTCAG GCCTTGCTCT GGTGCCAGGA GGAACTGATG CAGAAAAGAG 4801 TAAGAGGTCA TTTCCTGGAG GCTGTCACTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT 4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG 4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTTC TGGGTCTCAT 4981 CACCAGCTCC TGACCACCCG GTCAGCCCAT GTGCTTATTC CATAGCTTTC TTTTGCTATG 5041 TTTACTCAGT GTGGTGTTTG TTGGGACCCA GCAGAGCCA GTCCCAGGCT GACAGCTGTG 5101 GATACACAGG GCAGCATGAG GGTCCTCAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG 5161 ACCAGCACAC ATTCCTTCAA CCAACTATGT CTTGAAAAAC AAACATATTA TATCACATAT 5221 ATTGCATTTA TGAGACAGCT AAAATGTACT CGGGTAGCAT GACTCCAGGT GGGGATATCT 5281 GCAAGTGCCA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA 5341 CACAGCTTAG CTCCCTGGTG TTGGTTCAAA CTTTGAGGGT TTGACCACAA GCACTTTATT 5401 TTTGACATAT TTAAACAGAG CACAACTTTG GGAAAAAGTT TTCTTATGAA AATTATCACA 5461 ATAMAGCTTA AGGCATGACT ACATTAMAAT GCCTTTGCAA AGTATATGTG CCCTCTTCCA 5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT 5581 CAGGGATAAG TAAAATACTA AACTCTTTTG CAAAGTACAT AGACCCTCTT TCATAACAAT

Figure 5 (continued):

5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGGAA AGAGTCTAGC ATAAGCACGA 5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA 5761 CTACATTCAT TTCCAGTTTT CTGATCAGGC ACAGGTATGA ATCCCTTCTG TTGAAGAGAA 5821 AAGTCCATGT GTTTAAAATA TCTGGTTTCT CCAGTGCTAT TAGCGAGAAG ACTTGAGCCC 5881 TATACAACTC CCACCTGGAG TGACATCCTG TCTTCATGGT ATATTACATA CCTAGACACG 5941 CTCATCTCAC AGACTTAGGA CTTTGTCTTC TGATCTCCAT TTCTGATCCC ACTTCCACCT 6001 TTGCCTTGAT AGTGTCATTT TCTTCACTGC CTTGGTGACA ACCATGTTAT CCTCTGTGTA 6061 TTTGAGTGTT ACCATTTTCA GATTTTACCT GTATGCAAGA TCACACAGTC TTTGTCTTTC 6121 TGTCTGGATG CATGCTAATC TCTACACAAC AACCCTTCCC CGTCACTCAG ATCTTCCTCC 6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTCC CTTCAGTGGG GAGCTAGCTG 6241 GCTATTGGGC CTTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACTTAGAT 6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG 6361 ATCAGAGGTT CCAGTGTAAG ACACTTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG 6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCCTGGAAAC TGATGCAGGG CAAAGGACAG 6481 GTTATAAACC AAATCTTAGG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC 6541 TGCTTAGGGG CTACCAACCC AATCCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT 6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCCTGGG TCCAGCCTTT 6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGTCTTAG AGAGCCTAGG GATACTGGGT 6721 CCAGCCTTTG GTACCTTCAC TCTCAGGGTA CCCCATCACT GTGTCTTGGA GAGCCTAGGC 6781 ACCCTGGGTC CAGCCTTCAG TACCTGCGCT CTCAGGACAC CCCACCATTG TCTCTTGCCC 6841 OGTOTOTOT TOCTOTTOCT COCTITICATI GICTOTTOTO TGITTOTTIC TTGACTOTOC 6901 TITCCCCTCA CACCCTCACT CTAGITCTCC CCTTCCCTCT CTGCATCACC CTATTCTCTC 6961 TGTGGTCCCT CCACTTTCCT TTATCTCTCA TGCTTCTCT CTCCCTCAAA TACTTGTCAC 7021 CCACTATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC 7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAACTCTCT CCAGGTCCAA TTTCAGGTTC 7141 CTTCGAAAGC TTTCAGCAAA TGCTCAGGGA ACATGCCACT AACAAGAAGA TGCAAATTCC 7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCATCTTC CAGGCCAAGG TCAGAGGGGC 7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTTT GTTTTTAAGG TTTATTTATT 7321 TTAGGTGTTA GTGTCTTTGC TTGCATGACC TTATGTGCAT CATGTGTGTG CAGGTTCCTG 7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAC TGAGCCATCT 7501 CTCTAGGCCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTGTGTGTC AACATTGAGG 7561 AAAGGGAAAT AAAAAAAAA CTTCAGCCGC TAAGGTTGTA CAGTTTCACT AATTGCTACT 7621 TTTAGTTGTG ATAAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAAC TTCCCTGCTG 7681 GTGGTTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTTCTGT CTCTGTCTCT 7741 GTCTCTCCC TTCCATTCTT TCTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT 7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCCTTG GTGGGTGGTG TTACCAGCTG 7861 ATTTTTGGTT GTCCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGTCTCTG TCTGTATGAG 7921 TGAGGGAAGT CTGGGGAGCA GATTCCCTAA CCTTCAGCCT GGCCTGGTTC CTGAGTGAAC 7981 CCAGCCTCTC TGGTCCTAGT AGCTTTTTCC AAACAGGAAT CTGAGTGGTG ACAGGGAACA 8041 AGTACCAGCC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG 8101 ATTAGTAGTT GTATTGGATG TAGGAAGTCC TATCCTGGGA CAGCTAATCC TTAATGCTTC 8161 ACTGGAGATT TTCAATGAGA AATTTATCCC ACGGCCCATA TGGCCCCATC CTTTTGTCTC 8221 CAACAGCCAA GTATTITCCA TIAGAGGAGA CITCCIGTAC ACTIGATGGA TGCTCATTCC 8281 AAGGTGACTT GGGGCAGTCA GTACAGACTT GGGATGACCT CTGACAGCCT AACCTCTCCC 8341 CAACAAGGGC CCTCTATGTT TGCTATGTAA TGTAATGTCA GACATTGTCA GGAGTGTCCG 8401 CAGCACAGCC TGCCCAGTGT GAGGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG 8461 GATAACGAGG AGGTAAGCTG CAGTTCCCAG TCTCACTTCA CAGAGGAAGA GATAACCCCA 8521 TCCCAGGTCA TGTAGCCAGC AGTGGAAAGA ATGAGGATTT GAACTCAGGT CTTCCAAGTC 8581 CCATTGATAG CATCTCCTCA CAAGTCCCTT GCCACCCTCA CGATGCCTTA GACACTTGCC 8641 TGCCCTTTAT ACTAAGGAGA TGCAGGTACA AGGGGTTTAC CCATGTAGCA GCTGAGGCAG 8701 CTGGGGATAG ATACCAGCAG CAGGCCTGAT GTCACCACTC TAACTCCAGC ATCCCCAGTC 8761 TGTGTTCCTG GAGTGTGAAA ATCCCTACTT AACAAGATTG TGCAACAGTC CTTGGCTCTG 8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTTGTGGA ACATGGTGGC AGCCAGCCAA 8881 ARAGAGGGTC TGCATACAGA AGACACGTGT GGCAAGGCCA CAGCAGACTC TGACTACCTT 8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTTGG TCACCTCATG TTAAGGACAG 9001 GCCCTAATGA AGATGGGGCA GAAGACTGAA GGAATGGCCA ACCAATAACT GGCCCAACTT 9061 GAGACCCATC CTACAGGCAA GCATCAATTC CTGACACTAC TAATGATACT CTGTTATGCT 9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC 9181 AGAAAAAGAT ATCCACAGGC AAACAGTGGA TGGAGGTCAG GGACTATTAT GGGAGAGCTG 9241 TGGGAAGGAT TAAAAACCCT GAAGGGGATA GGAACCCCAC AGGAAGACCA ACAGAGTCAA 9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCAA CCAAAGAGCA TACACAGGCC 9361 GGTCCGAGGC ACCTGGCACG TGTGAAGCAG ACATGCAGCT CAGTCTCCAT GTAGGTCCTC 9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCCAACA GGGCTGCATA GTCTGGCCTC 9481 AGTGGGGGAG GATGCCCCTA ATCCTGCAGA GACTTGATGA GTGGAGAGCT ATCCAGGGGG

Figure 5 (continued):

9541 AACCCACCCT CTCTGAGAAG GGAATGGGGA TGGGGGAGGG ACTCTGTGAA GAGGGGACAA 9601 GGACAAACAA GAACCTCAAA TAGGTCAGGC CCTAAAGGCT TGCTAAGTAG CAGTGGCCCA 9661 GCTCTGTCCT GTTCCTCAGC CCAAGGCTCA GCTCCCACCT GTTTCTGTGT TTTTCTGGCT 9721 TTTCATGGGC CTAGGACTTG GTGACCAGTT CAAACAATGG GGCCTGTGGA AGACACAATA 9781 TACAAGACTA GGGACATTCC TGTTCTGCTG ACTATCCATA GCCTGATGTA GGTGGAAGGA 9841 CCCAATCACT GGATTTCTAC CCTTGCACAA CCTTGACAGC TGAGGGCCTC TCAGAAACCT 9901 ATTTCTCCA CTGAAAAATG AGACTCTCAA ATGAACGTCG TGACAATCAT CAGGCTTATT 9961 ARAGAGGTGT ATCTARCCTG AATGGCAAGC AGACAGCAGG CARATGTCTG TATCAACCTC 10021 TAGGAAGGAC AAGAACTGCT CACTGCTGCC CCCCAGGAGG CCATTTGCTG AAACAGCTGC 10081 TCTCCTGCTG GTGCACAGGC CCTGCCTTCT CATTGCAGCC ACAGCCCCTT CCTGTCTGAA 10141 CCTCCTGTCA GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAAA 10201 TAAAAAACAG TGGTCATGGC TATTCATAGG GGTTTATGCT TCTTCAGTCC ACACTGTGAA 10261 GAGCTGTGGG CATGAACCAC AGTGTTCGAG GTAGAGTTGG GGTTCTGAAA TTCACAGTGG 10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAGA CACACAGTCC TGCTGCTTCT 10381 GITCCCAATA TCCTGAGGAG ACGACACCT TACTTTGTTC AGAGGCCACA GTCTAGTTGA 10441 CCTGAGAGTT ACCAGTTTCT TATTTGTGTG TGTGTGTGT TGTGTGTGTG 10501 TGTTGTTCGT GTGTGAGTGC AGGTGCACAT ATGATAGCGT ACACGTTGAG GTCAGAGGAT 10561 AACTATCAGG CGTTGTCCCC TCCTACTTTT CCTCGGACTC TGGAGAACAA ACATGGGTCC 10621 TTATTCCAGG GGAGCAAGTC GCTGTTGGCT GACACATCIT GCTCACATAC ATTTTACCTA 10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCCTCA CCGATGGCAA TGCACCACCT 10741 CTCTACCCAC ATAGGAGTTG GGTCTCCACA CACCCCCACA CCCCCTTCAC CAAAACGTTT 10801 TCAGTTACTT TATCTGGTAA AGTTCATCAG AGAATGAAGC CAGTATTAAG AACATGGAAT 10861 CATTTGGGAA CCTGGATCTA GCAATACCCC ACCCTAGATG GAGTTGCTGA GTTTTCACCT 10921 CAGATTATAA TTCCCCCCTA GCTTCTATGG TTTATTCTGA AACCAGGGGA ACTCGATTCC 10981 TCCCTTTGGA CCACAGACAT CCTGGCTTGT GAATTCACAT GTCATCTACT GCTAATCCAT 11041 TGGTAGTATG TGGCTCACAG AGACACATA CAGTCATGGC CAATGTCAAG GTAGGACAGA 11101 TGTGAATCAT TCCCCCAGTC CTGCTGTTTT CATGACTAAC CCTCCTCAGC ACAGTGACCA 11161 TGAACCTACT TTTCCCCTCC TTTTATTTTT AGAATTGCTG GAATTTTCTA TTTTGAGAAA 11221 TANTAGCCTT GGGCAGCATT AAACAAAATC ATCTAGAAAG CTGGTTTAAA ATACAGATGG 11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCAT TATTGGCCCC TCACAGAGGC TGGCTCACTC 11341 CAGCAGAGGT GGTTGAAGCT CTTGGACACG GGTCAGGTGC ATAGGAAAGG TNGTCTGGGA 11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTTGGCTTT TTTTTTTTTA AATGAGTTCT 11461 CAAAAATGA CTGGCTAGCT TAGGCAAATA CTTCGAGCCA ACCCAACAGA ACATTCTTCC 11521 ATTGATTCAT TCTGGATCIT CTTTCTAGAC AATACTGAAC TGACCCCTTG TTGGCAGTCT 11581 CAAGTTTGAC AACATAGGGC TTTGAACTTG GCACAAGGTC CATCACTGTC ACCCAAGCAT 11641 CCTGGGTGAC CTTTGGGTTG GAATATCTTG GCTAACCTTA GATATTTTCT TTGGAGTATC 11701 TITAGAACAT CCAGGAAATA GGGCTTGATT CTCATCCTGG GACCACAATA TAAGTCACCC 11761 TAGAATCCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCAT CCTTCTTCAA 11821 AGCAGTGAGT AGTGACTCCA CTAAACTGAG TTCCCATCTG AGAGTCCACA GGAGGCTTTG 11881 GGGCAAGAAG CAGAGGGAAG GCACTGITTG TGTTGGTAAA GTTTTGACTC TAACAAATTT 11941 GAAGACATAG ATGACATTGT GTCAGACTAA CAACAACCTA GACTCATGTG GGTTCTGTTT 12001 AGGGATCAGA TITTATTCAT CAATGACTTG TCTTAGTGTA TAGAGAAAGG CTTCCTACTG 12061 GAGTGTAGGC TCAATAATGA CAGAAGAGAT AGCTATTTCC CCTAGGGACT GTGCTGCTCC 12121 AAGITTGGTG GAGAAAGGCA GTGGGGAACC TAGATGTGCT CTCTGGGGAG GGGGTCTGAA 12181 GCTGGCTTCA TAGAAGGTGT GAAGTTTTGC TGAAACATCT AAACAGAATT ATAGCTTAGG 12241 AAAGTGAGCA GGCAAGGCAG GGAATGTGTT GCATATGTAT ATGTACATGA ATATATTATG 12301 TTATAGATAC ACACACATTI GAACCTCATI TGCAGATGAC AGAAAATAGG TTATTTTGCC 12361 TCTCTTAACT GCTAAGCACA ATGACTTCCA GTTCCATCCA TTTCCTGAAA TGCCACAATT 12421 TCATTTTCA TTGTGGCTGA ATAAAATTCC ATTGCAGACT GGGCCCTACT TCATCCACTC 12481 CTGAGGGCAG GCATATCCCC TGGCTCCATT TCTTACCTAT TGTGAAGAGA AGTGCAACTG 12541 TCTTGTTGAA AGGCAAGCGT GAGAGAGGCA GGCACTAATT GTGGGTTTTT GTTTCTTCTT 12601 CCTGCTATGA CTCTCCATTT GTCAGAACCA AAGATCGATA AAAGCCGCCA CCATGAAAGC 12661 CATCTIAATC CCATTITTAT CTCTTCTGAT TCCGTTAACC CCGCAATCTG CATTCGCTCA 12721 GAGTGAGCCG GAGCTGAAGC TGGAAAGTGT GGTGATTGTC AGTCGTCATG GTGTGCGTGC 12781 TCCAACCAAG GCCACGCAAC TGATGCAGGA TGTCACCCCA GACGCATGGC CAACCTGGĆC 12841 GGTAAAACTG GGTTGGCTGA CACCGCGCGG TGGTGAGCTA ATCGCCTATC TCGGACATTA 12901 CCAACGCCAG CGTCTGGTAG CCGACGGATT GCTGGCGAAA AAGGGCTGCC CGCAGTCTGG 12961 TCAGGTCGCG ATTATTGCTG ATGTCGACGA GCGTACCCGT AAAACAGGCG AAGCCTTCGC 13021 CGCCGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGCAGATA CGTCCAGTCC 13081 CGATCCGITA TTIAATCCTC TAAAAACTGG CGTTTGCCAA CTGGATAACG CGAACGTGAC 13141 TGACGCGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTTACCGGGC ATCGGCAAAC 13201 GGCGTTTCGC GAACTGGAAC GGGTGCTTAA TTTTCCGCAA TCAAACTTGT GCCTTAAACG 13261 TGAGAAACAG GACGAAAGCT GTTCATTAAC GCAGGCATTA CCATCGGAAC TCAAGGTGAG 13321 CGCCGACAAT GTCTCATTAA CCGGTGCGGT AAGCCTCGCA TCAATGCTGA CGGAGATATT 13381 TCTCCTGCAA CAAGCACAGG GAATGCCGGA GCCGGGGTGG GGAAGGATCA CCGATTCACA

Figure 5 (continued):

13441	CCAGTGGAAC	ACCTTGCTAA	GTTTGCATAA	CCCCAPTT	The desirement of the Co	AACGCACGCC
13501	AGAGGTTGCC	CGCAGCCGCG	CCACCCCGTT	ATTAGATITG	ATCARCACAC	CGTTGACGCC
13561	CCATCCACCG	CAAAAACAGG	CGTATGGTGT	GACATTACCC	NI CHURCHCHO	TGTTTATCGC
13621	CGGACACGAT	ACTAATCTGG	CAAATCTCGG	CCCCCACTC	ACTICAGIGE	GGACGCTTCC
13681	CGGTCAGCCG	GATAACACGC	CCCAGGTGG	TCACOCACTO	TTTC A A COM	GGCGTCGGCT
13741	AAGCGATAAC	AGCCAGTGGA	TTCACCTTTC	CCTCCTCTC	TITGAACGCT	AGCAGATGCG
13801	TGATAAAACG	CCGCTGTCAT	TARATACCC	CCCCCCACAC	CAGACITIAC	AGCAGATGCG CCCTGGCAGG
13861	ATGTGAAGAG	CCADATCCCC	ACCCCATOCCC	TTYCETYCOL	GIGAAACIGA	AAATCGTGAA
13921	TGAAGCACGC	ATACCCCCTT	VOCOCYTOTO	11CG11GGCA	GGTTTTACGC	AAATCGTGAA
13981	AAGAGGAAGA	ACAGAAGGAT	CCABILIGIA	AGGIALLUGG	GGATCACAAC	TTGCCCTCTG
14041	TTACTTCTCA	ACTOM TOOM!	CCACACIC	TCCTGCTGGC	TACTCTCCAG	TEGTTTCATC
14101	GACCACCCAA	ACCACATITE	CICIAGAAAG	TGCTACTATC	ATCCACACAT	TTCTACCTGA
14161	CACCATCCCA	CAATTAAAAT	CAAATICICI	TCCTCTCTGA	GTAGTCTCCA	CACCTGTTAC
14221	CACCATCCCA	CTTCCCACTC	CCTAACTGCA	CTCTGGCGTG	TGACTTGCCT	CAGTCCTTGC
14281	AATAAGAGTT	GIIGGCAGIG	CUAGGCGIGG	TGGCGCACGC	CTTTAATTCC	AGCACTTGGG
14241	AGGCAGAGGC	AGGGGATTT	CIGAGITCGA	GGCCAGCCTG	GTCTACAGAG	TGAGTTCCAG
14401	GACAGCCAGG	GCIATACAGA	GAAACCCTGT	GTCGAAAAAC	САЛАЛАЛАЛА	AAAAAAGTT
14461	GTTGGCAGAG	TGTGGGTTAT	ATACCAGGTG	GAGATTTCAA	ATGAGTGGCT	GAAGCTGTAG
14501	CCAGAAGGAA	CTTAGAGGAT	AGCTCATAAC	TTAAAAAGAA	ATGTAGAGAG	TAGCAGAAAC
14501	ATTGAGAGAG	IGGGCACACA	GCCACTGTGT	GAATGIGGCA	GAACACAATC	CAGCCAGCTA
14561	TACATGCATA	AGIGIATATT	GGCGCCATCC	TGACTGATGA	GACACAGGAA	AACAGATAGA
14701	CGGGGTTAGG	TGGCCATGGC	CITTCCTGCC	TGCCTCTTCC	TAAGGGTCAT	CTCAAGACCT
14701	TATGCTCTCT	TAACICTICC	ATTGCTACTT	AGCTTCTAGA	TATCACCTCC	AGATTAGTCT
14/01	CCTTGGGTAC	ATCAGTGATC	CTGGTGATAT	CCAGGGCTTC	CTGATTCCAT	CTTTGTCATA
1402L	GAGGCTGCAA	CTAAAGAGGT	CTTCTTAATA	CTTCACACCC	TGATGCCAAA	AGGAAGACAC
14881	AGAAGTTCAC	AGAGGTGAAG	TGATTCATGT	AGGACATACA	GTGAGCAAGC	ATCAGGGTCC
14341	GGATTATCTG	ACTOTACTOT	AACTTTTATG	TAAATGTGCT	TTATGCCATT	AACACTGTCA
12001	TTCCTGTGCT	TCAGCTCTGG	GAGACTCCCA	AGCACTCTTA	GGCACAAGCC	ACAATTAAGG
12001	GACTCTGACA	CTCTGCATTG	ATTAATTAGC	ATGGTGGTCT	CTATGTTTCC	AGATTCATGA
15101	TTGTTTCACT	TICCATATAG	GCTATGAAGG	GTGTGAGGAA	ATTTTTTGGG	GACAGAATTG
12241	GAGGCAATCC	ACCICICICA	GGAAGCCTCT	ATCTGGAAAA	GCTTACAACT	CAGGGACAGT
15301	AACTGTAGGC	CLAGICCITG	GIGICCAAAA	TGGGTTTTAT	GGTTTGAATC	TGCAAAGCCT
15361	TCCATGTGCT	CAMAGGITIG	AACATGGAGC	CTCCTCCTGG	TAACACTGTA	TTGGAGGCTT
15421	TTGAGACTGG	AIGCICITIG	GICCCATGIT	TTGCTACATC	ATCTGTCAAG	ATATGACCCA
15481	GGCATGCTAC	TOTOGRADUA	AGACTATGCC	TCTCCAGCTT	TCATGTTCTC	CCCACCATGA
15541	TAGACTTGTA	CTCCCTAAAA	AIGGAATCAA	AGCAAACTTT	TCCTGCATTA	AGTTTTTTT
15601	TTTCTGTTAA	GIGITIGGIC	ACAGGGACAA	GAAAACACTC	AATACAGATA	ATTAGTACCA
15661	GAGTTGAGGT	TOTAGE	AGCAAGTTGG	ATCAAATTIT	TAGGGCTTTG	GAACTGATTT
15721	ATAAGAGACA AATAGTTTAA	TOLMOMBUMU	Champonera	TGGGCTACAG	AAGTGTCACC	AGTTTTTAAG
15781	AATAGTTTAA AAAACGTGAG	CATCONICCALGG	CMATTGTGAA	AATCAGAATG	CICACACAAA	GGCAGACAGG
15841	GAAGCCATTC	CAIGIGGCGI	GIGAGAGGGC	ATAAGAAGGA	ACCTAGGGGG	AAATGAGCTA
15901	CTGAATGAGG	CCBARTTTA	A CCA CTCCA	GIGGCIGIGC	TIGGCCCATG	CCCTGGCAAT
15961	CAGACCACCA	CCWWIIII	WORNETOCH	CTAACTOGAT	TGTCAGAGAA	AATATCAAGA
16021	CAGACCACCA TTGTGAAATT	CLCUDGCIVI	GCCG1G111G	TGACCGACCA	GCTACTCTTA	GCCAGCTCTA
16081	GGTGTGGGTC	CCMGMGCMAI	TATCAGAGCA	TGAAGATACA	TACAGTTTAG	TGAAGTAAGG
16741	GATAATCCAA	ADTATOROGA	10G1GCAIAA	ATCTATGTAG	GTGATGCCTA	AGTGACACTT
16201	GATAATCCAA TAGAACTTTG	WINICHOCK	MIGIOGAMIG	TCTTCCAAGG	AGACCTGTAG	ACACACATTT
16261	GTCTGAGTTA	CICATOCCIO	INVINACING	CIAGCIAGAA	ATCATTTCCT	GAAGAGGTTA
16321	AGCCAAAGCT	COCTTCCTCC	GCMATCHIT.	AGTGATGGCA	AGGAAGGCAT	TGCAGTCAGG
16381	AGCCAAAGGT	TTATABAACC	CALIGUATUA	AGAGTAGAGA	GTCAGAGTGT	Gagtagaaag
16441	AGGATACAGG	TIMINAMACC	COCCUCA	CTCTCAGCAA	TCCATTTTCT	CCTAAAAGGC
16501	TTTACCTTCT	WWWWTTIIW	GICTICAAAA	CCAGTACCAG	TAGCCTGGGA	ACAAAAGTTG
16561	AAACAAATGA	CTROCKRO	GGCATTTCAC	ACTTAAAACA	GGGCATCACC	TAGGAGGAGC
16621	CCTGTGTGCA	GINGGANGIG	TOGCCTCTGT	GTCAGGAATG	CTCAGGCTAA	TAAGGGGTCC
16601	TCTATCTGAG	GGACCCTATG	AAGATTCAAC	AAGTAGTTGT	GAGAATTCCC	TGTAAATGGA
16741	TGCTACCAAT	TIGACATTIG	TAGACCTGCT	ATTGTGTGCT	TCTTTATTGG	GCTCTCCCAT
16801	CTCCCAACTT	CARCCUAT	ATTUCACATT	AATCCCTTCC	ACCACCATGC	AACACTAGGT
16061	AGGAGAGAAG	GAAGGTTAGA	AGAGAAAGTG	GGTATAGATC	TATTTAGACT	ACTICCIGCT
70001	GATTAGGGGC	AAGTCCAATC	GICATTGTCA	GGATACCTCC	AACCAGCAAC	CAGCAAACCA
10371	GCAAATCAGA	AACAGCAAAA	GCAGCCAACA	AGGCAGCACT	AACCACCACC	N TOTAL CONTROL
TOART	GTAGCGTGGG	AGCAGTCACT	ACTGGTCTTC	TCATGGGTTT	ርርር ግጥ ተለ አጥአ	CALCALCANCES & C
1/041	AAATTCCGTA	ATTITITCCC.	CACCACCTGA	AATTOTTAA	TTTTNNNTCC	8 8 8 CM 8 MCC 8
11101	CAGCIGGCAA	AAATCACATC	TCTCCTAGAG	CACAAGACAA	ልጥሮክ ምክርምምክ ፡	
T \ T P T	GCAATCTGAA	GCATCTCAAT	ATCCCACACC	TGGGATTAAA	המשמממשת	תיישיים ביים מיישים
1/221	CATAACTGTT	TTTTTTTTCC	TATTTTTAT	TAGGTATTTT	שמידית מידידים כי	THE A ANTONIA
1/281	CTATCCCGAA	AGTCCCCTAT	ACCCTCCCAC	CICCCICCIC	CCCTACACAC	CCACTCCCAC
			15/5			_
				-		

Figure 5 (continued):

17341 TTTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCCT 17401 CTCTTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG 17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTCA 17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGCTC TGTGTTTTAT CTAATAGATG 17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA 17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT 17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GCGGCCGCCA CCGCGGTGGA GCTCCAGCTT 17761 TTGTTCCCTT TAGTGAGGGT TAATTGCGCG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC 17821 TGTGTGAAAT TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG 17881 TARAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC 17941 CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG 18001 GAGAGGCGGT TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC 18061 GGTCGTTCGG CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC 18121 AGAATCAGGG GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA 18181 CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA 18241 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC 18301 GTTTCCCCCT GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA 18361 CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA 18421 TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA 18481 GCCCGACCGC TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA 18541 CITATOGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG 18601 TGCTACAGAG TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTTGG 18661 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG 18721 CAAACAAACC ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG 18781 AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA 18841 CGAAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT 18901 CCITITAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC 18961 TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC 19021 ATCCATAGIT GCCTGACTCC COGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC 19081 TEGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC 19141 AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC 19201 CATCCAGTCT ATTAXTTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT 19261 GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC 19321 TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA 19381 AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTCG CCGCAGTGTT 19441 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG 19501 CTTTTCTGIG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC 19561 GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA 19621 AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT 19681 GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT 19741 CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG 19801 GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA 19861 TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT 19921 AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTAAA TTGTAAGCGT TAATATTTTG 19981 TEAAAATTCG CGTTAAATTT TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC 20041 GCCAAAATCC CITATAAATC AAAAGAATAG ACCGAGATAG GGTTGAGTGT TGTTCCAGTT 20101 TGGAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC 20161 TATCAGGGCG ATGGCCCACT ACGTGAACCA TCACCCTAAT CAAGTTTTTT GGGGTCGAGG 20221 TGCCGTAAAG CACTAAATCG GAACCCTAAA GGGAGCCCCC GATTTAGAGC TTGACGGGGA 20281 AAGCCGGCGA ACCTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG 20341 CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG 20401 CTACAGGGCG CGTCCCATTC GCCATTCAGG CTGCGCAACT GTTGGGAAGG GCGATCGGTG 20461 CGGGCCTCTT CGCTATTACG CCAGCTGGCG AAAGGGGGGAT GTGCTGCAAG GCGATTAAGT 20521 TGGGTAACGC CAGGGTTTTC CCAGTCACGA CGTTGTAAAA CGACGGCCAG TGAGCGCGCG 20581 TAATACGACT CACTATAGGG CGAATTGGGT ACCGGGCCCC CCC

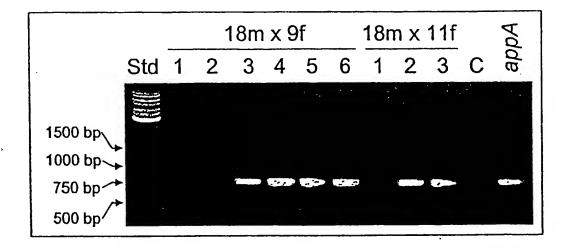


Figure 6

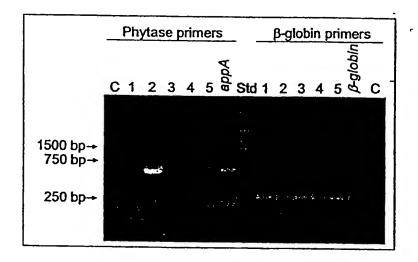


Figure 7

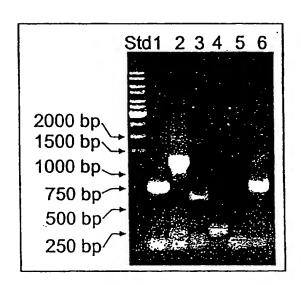


Figure 8

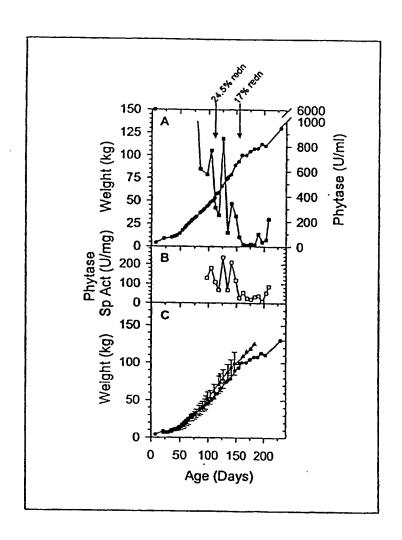


Figure 9

20/58

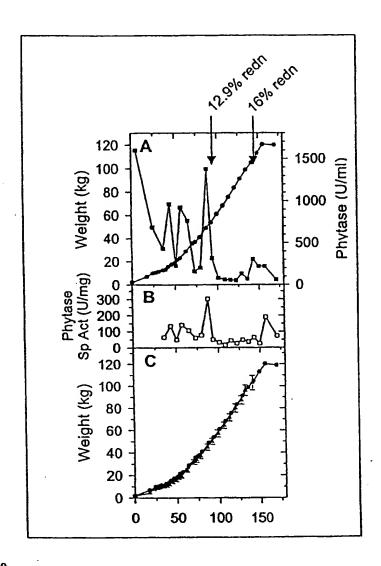


Figure 10

WO 00/64247

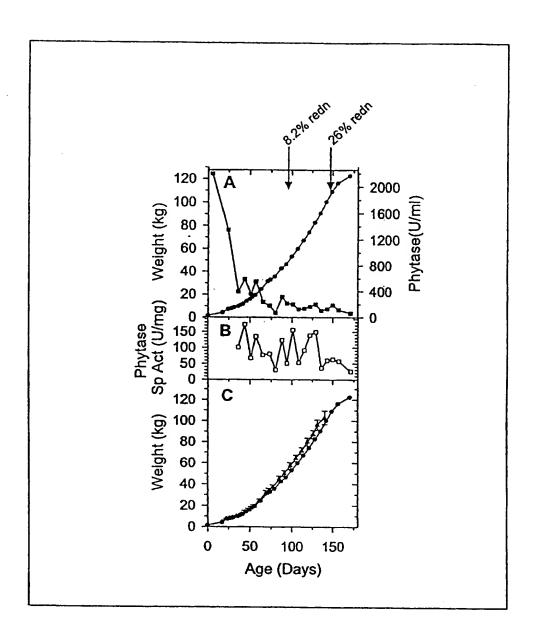


Figure 11

22/58
RECTIFIED SHEET (RULE 91)

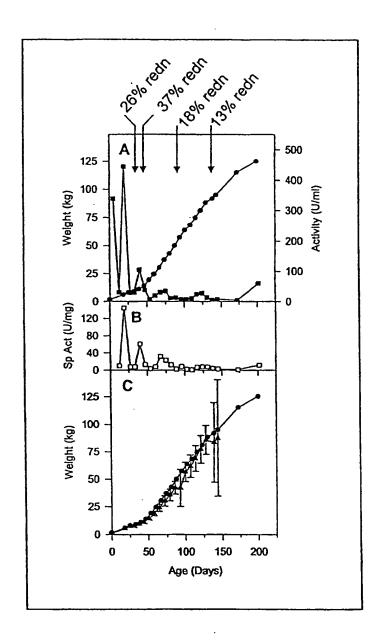


Figure 12

WO 00/64247

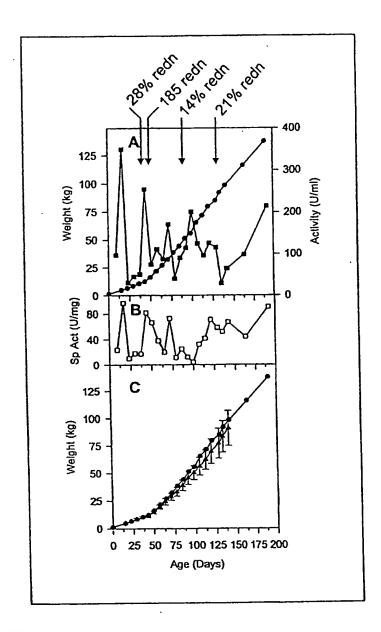


Figure 13

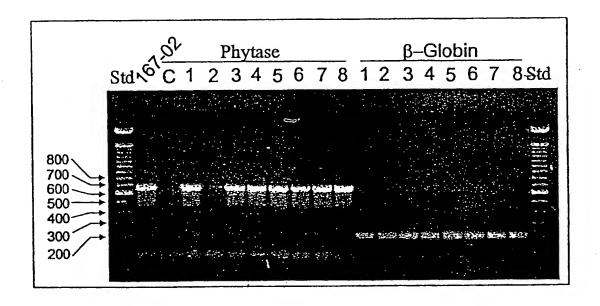


Figure 14

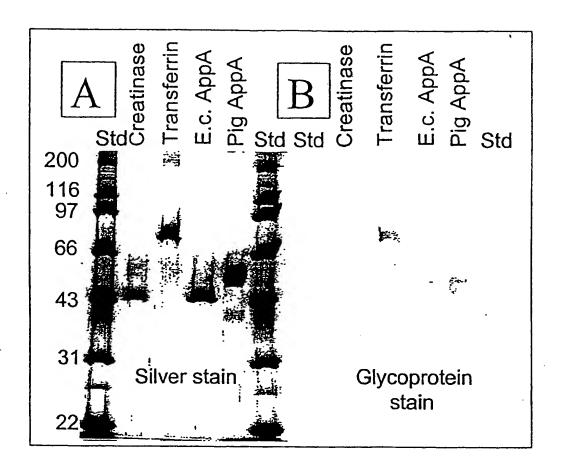


Figure 15

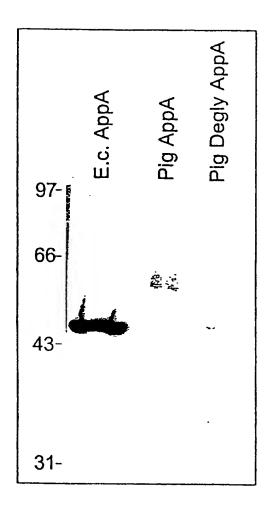


Figure 15B

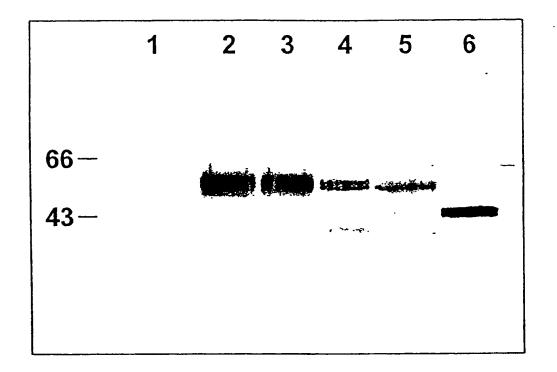


Figure 16

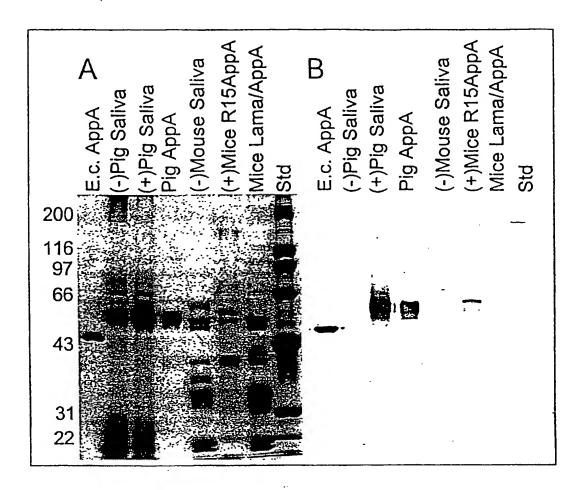


Figure 17

Figure 18: Nucleic acid sequence of the known segment of the R15/appa+intron plasmid, including the vector sequences of pBLCAT3 (SEQ ID NO:2).

```
LOCUS
                               6708 bp
            R15/appa+intron
                                         DNA
                                                 SYN
                                                            15-APR-2000
DEFINITION R15/appa+intron transgene with vector cut 13543 to 4954
ACCESSION R15/appa+intron
REFERENCE
           1 (bases 1 to 6708))
SOURCE
            synthetic construct.
      ORGANISM synthetic construct
                 artificial sequence.
KEYWORDS
            salivary proline-rich protein, acid glucose-1-phosphatase; appA
            gene; periplasmic phosphoanhydride phosphohydrolase; artificial
            sequence;
AUTHORS
            Golovan, S., Forsberg, C.W., Phillips, J.
  JOURNAL
            Unpublished.
    DEFINITION Rat salivary proline-rich protein (RP15) gene.
    ACCESSION
                M64793 M36414
    VERSION
                M64793.1 GI:206711
                Rat (Sprague-Dawley) liver DNA.
    SOURCE
      ORGANISM Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
                Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
    REFERENCE
                1 (bases 1 to 1748)
       AUTHORS
                Lin, H.H. and Ann, D.K.
                Molecular characterization of rat multigene family
       TITLE
encoding
                 proline-rich proteins
                Genomics 10, 102-113 (1991)
       JOURNAL
       MEDLINE
                91257817
     FEATURES
                          Location/Qualifiers
          source
                          1..1748
                          /organism="Rattus norvegicus"
                          /strain="Sprague-Dawley"
                          /db xref="taxon:10116"
                          /tissue_type="liver"
                          /tissue_lib="cosmid genomic library"
           misc_feature
                                 1802-1810
                          /function=" consensus sequence for initiation in
                                      higher eukaryotes *
                     Location/Qualifiers
DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
```

FEATURES gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375

VERSION M58708.1 GI:145283

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision;

Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 1811..3109)

> AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 18 (continued):

TITLE The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

MEDLINE 90368616

FEATURES

Location/Qualifiers 1811..3109

Source

/organism="Escherichia coli"

/db_xref="taxon:562"

sig_peptide

1811.. 1876

/gene="appA" .

ന്നട

1811..3109 /gene="appA"

/standard_name="acid phosphatase/phytase"

/transl_table=11

/product="periplasmic phosphoanhydride

phosphohydrolase"

/protein_id="AAA72086.1" /db_xref="GI:145285"

translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP/

TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPOS

GQVAI IADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLPNPLKTGVCQLDNA

NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS

ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF

YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG

ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat_peptide

1877 3106 /gene="appA"

/product="periplasmic phosphoanhydride

phosphohydrolase*

mutation replace(1817.. 1819, "gcg changed to gcc")

/gene="appA"

/standard_name="A3 mutant"

/note="created by site directed mutagenesis"

/phenotype="silent mutation"

mutation

replace (3092..3094, " ccg changed to ccc")

/gene="appA"

/standard_name=" P428 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

mutation

replace(3095..3097, gcg changed to gct*) /gene="appA"

/standard name=" A429 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

Figure 18 (continued):

```
DEFINITION Plasmid pBLCAT3 (bases 3109 to 6708)
    ACCESSION
                X64409
    VERSION
                X64409.1 GI:58163
    SOURCE
                synthetic construct.
      ORGANISM synthetic construct
                artificial sequence.
                1 (bases 3109 to 6708)
    REFERENCE
                Luckow, B.H.R.
      AUTHORS
      TITLE
                Direct Submission
                Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
      JOURNAL
                 Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
                2 (bases 3109 to 6708)
     REFERENCE
                 Luckow, B. and Schutz, G.
       AUTHORS
                 CAT constructions with multiple unique restriction sites
       TITLE
for
                 the functional analysis of eukaryotic promoters and
regulatory
                 elements
       JOURNAL
                 Nucleic Acids Res. 15 (13), 5490 (1987)
       MEDLINE
                 87260024
                 Promoterless CAT vector for transient transfection
     COMMENT
experiments
                  with eukaryotic cells. Allows the analysis of foreign
                  promoters and enhancers.
                          Location/Qualifiers
     FEATURES
                          3109 to 6116
          source
                          /organism="synthetic construct"
                          /db xref="taxon:32630"
                          3197..3810
          SV40 t intron
                           /note="SV40 signals"
                           3807..4047
          polyA_signal
                           /note="SV40 signals"
                           complement (5244..6104)
          CDS
                           /codon_start=1
                           /transl_table=11
                           /qene="Amp"
                           /product="beta-lactamase"
                           /protein id="CAA45753.1"
                           /db xref="GI:58165"
 BASE COUNT
                                  1515 g
                                           1798 t
                1916 a
                        1479 C
 ORIGIN
         1 GGATCCCCTT TGCTATGTAG TTTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
        61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
       121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
       181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
       241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
       301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
       361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC
       421 AGGTCAACAG TGCCACATAT CCITTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
       481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
       541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
       601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAA TAGACAAACA TAGTTAATTG
       661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
       721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
       781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
       841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
                                     32/58
```

Figure 18 (continued):

	TAAGATAAAG					
	TTCAGCTCTA					
1021	TCCATGGACC	TTTGAAATAT	AAAATAGTCA	AGCAACTTAT	CAAGGAATTA	CAGATTCCTT
1081	GATACTAACA	CAGGTAAATC	CCACACGTGT	TTTGAGACTA	CATTTGCTGG	GATTTTATTG
1141	ATGTAATAGG	TCACATGTTT	TTCGGGCCAA	TGTTGCTGTT	ATTCGGTTAC	TTCAAGAGAA
	TAGTGGCAAC					
	GTAAAAGAAT					
1321	GTGTTTAAGC	TGTACTATTG	ATCAAAGAAA	TTTATTACCT	TCAGTTTCAA	TGGAAATAAT
	TACTGATAAT					
1441	CAGAAAATAT	TAGCAAGTAG	AATGCAATAT	TTATATAACG	ATTGTATTTA	TCAATCAATT
1501	GTATGTATCA	ATATATGGGC	TATTTTCTTA	CACATGATTT	TATTCAAATT	TACTCTAATC
	ATTGTTGAAC					
1621	AAAAGTCCCA	GTGTGGAGTA	AAGGATGCAA	GATTTCCTGC	TCTGTTAAGT	TAAAATAAT
	AGTATGAATT					
	CCAGCACAGA					
	AGCCGCCACC					
	GCAATCTGCA					
1921	TCGTCATGGT	CACCACACA	CAACCAAGGC	CACGCAACTG	ATGCAGGATG	TCACCCCAGA
1981	CGCATGGCCA	ACCTGGCCGG	TAAAACTGGG	TIGGCTGACA	CCGCGCGGTG	GTGAGCTAAT
	CGCCTATCTC					
	GGGCTGCCCG					
	AACAGGCGAA					
	GGCAGATACG					
	GGATAACGCG					
	TACCGGGCAT					
	AAACTTGTGC					
	ATCGGAACTC					
	AATGCTGACG					
	AAGGATCACC					
2641	TITGCTACAA	CGCACGCCAG	AGGTTGCCCG	CAGCCGCGCC	ACCCCGTTAT	TAGATTTGAT
2701	CAAGACAGCG	TTGACGCCCC	ATCCACCGCA	AAAACAGGCG	TATGGTGTGA	CATTACCCAC
	TTCAGTGCTG					
	GCTCAACTGG					
	TGAACGCTGG					
	GACTITACAG					
	GAAACTGACC					
	TTTTACGCAA					
	GTTATTGGTG					
	TGGCAGAAAT					
	ACAAACTACC					
3301	GIGTTAAACT	ACTEATTOTA	ATTICTUTES	TATITTAGAT	TCCAACCTAT	GGAACTGATG
3361	AATGGGAGCA	CTCCTCCDAT	CCCTTTAATC	AGGAAAACCT	GTTTTGCTCA	GAAGAAATGC
	CATCTAGTGA					
3481	GAAAGGTAGA	AGACCCCAAG	GACTUTCCT	CAGAATTGCT	AAGTTTTTTG	AGTCATGCTG
3541	TGTTTAGTAA	דרדים מבומת . דידים מבומת ה	GC-LAGG-LALAG	CTATTTACAC	CACAAAGGAA	AAAGCTGCAC
3601	TGCTATACAA	מדמדממממם	דרמדממממם:	CTGTAACCTT	TATAAGTAGG	CATAACAGTT
3661	משתאת מידה . ממידמים מידה	ראדארדורוט	י דיייאריייייייייייייייייייייייייייייייי	CACACAGGCA	TAGAGTGTCT	GCTATTAATA
2721	A COUNTY COUCH	רביים ביים אונה ביים אונה ביים היים היים היים היים היים היים היי	י אריים אודים דיים כריים	יייייייי איייייייייייייייייייייייייייי	TAAAGGGGTT	AATAAGGAAT
3721	ACTAIGCICA	TARRETTOIGE	ACCITIAGE:	י אייאאייאמרים	י אדאררארארא י	TGTAGAGGTT
2013	THE STATES	TWGTOCCTIC	· CCCPCYCCCC	י ההההשמשה. י שושעורשפרר	ממדמסממדוי	AATGAATGCA
204	LIMCTIGCTI	TARRAMANCCI	, CCCMCMCCIC	ייייייייייייייייייייייייייייייייייייי	מתמשת משום	CAATAGCATC
390	ALIGITOTIC	LIMMCTIGII	Mulalatatatata	/ CALCINGAICADD	Ghildhithtana Teannar	GTCCAAACTC
396	ACAAATITCA	CAAATAAAGC	ATTITITU	r CIGCHIICIA	. Gligicalcus.	TTCGTAATCA
402	ATCAATGTAT	CTTATCATG	CIGGAICGA	, CCCCGGGTWC	, COMUCICUMA	CADCAMATCA
408	I TGGTCATAGC	TGTTTCCTGT	GIGAAATIG	TATUUGUTUA	TOWN TO CHURCH	CAACATACGA
414	L GCCGGAAGCA	TAAAGIGTAA	A AGCCTGGGG	CONNECTION	A TOMOCIANCI	CACATTAATT
420	GCGTTGCGCT	CACIGCCCGC	TITCCAGTC	J GGAAACCIGI	COTOCCHOCI	GCATTAATGA
426	ATCGGCCAAC	CUCUCGGGA	AGGCGGTTT	CGTATTGGGC		TTCCTCGCTC
432	l ACTGACTCGC	: TGCGCTCGG1	CGITCGGCT	s CGGCGAGCG0	TATUAGUTUA	CTCAAAGGCG

Figure 18 (continued):

11

```
4381 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
4441 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
4501 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
4561 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
4621 CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA
4681 TGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG
4741 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC
4801 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
4861 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
4921 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
4981 GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
5041 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
5101 TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
5161 AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA
5221 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
5281 ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
5341 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
5401 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
5461 GCAACTITAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT
5521 TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
5581 TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA
5641 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT
5701 AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC
5761 ATGCCATCCG TAAGATGCIT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
5821 TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA
5881 CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA
5941 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6001 TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
6061 GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA
6121 TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT
6181 TAGAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
6241 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTTT
6301 CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG
6361 GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG
6421 GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
6481 GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA CCGCATCAGG
6541 CGCCATTCGC CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTTCG
6601 CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA
6661 GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGCTT
```

Figure 19: Nucleic acid sequence of the known segment of the R15/appa+intron transgene used for the generation of transgenic mice (SEQ ID NO: 3).

15-APR-2000 SYN 4060 bp LOCUS R15/appa DEFINITION R15/appa transgene without vector ACCESSION R15/appa REFERENCE 1 (bases 1 to 4060) synthetic construct. SOURCE ORGANISM synthetic construct artificial sequence. salivary proline-rich protein, acid glucose-1-phosphatase; appA KEYWORDS gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence; Golovan, S., Forsberg, C.W., Phillips, J. AUTHORS JOURNAL Unpublished. DEFINITION Rat salivary proline-rich protein (RP15) gene. ACCESSION M64793 M36414 M64793.1 GI:206711 VERSION Rat (Sprague-Dawley) liver DNA. SOURCE ORGANISM Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 1748) REFERENCE Lin, H.H. and Ann, D.K. AUTHORS Molecular characterization of rat multigene family TITLE encoding proline-rich proteins Genomics 10, 102-113 (1991) JOURNAL 91257817 MEDLINE Location/Qualifiers FEATURES 1..1748 source /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" /tissue_type="liver" /tissue_lib="cosmid genomic library" misc_feature 1802-1810 /function=" consensus sequence for initiation in

PEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)

gene,

higher eukaryotes *

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375

VERSION M58708.1 GI:145283

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision;

Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1811..3109)

AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 19 (continued):

The complete nucleotide sequence of the Escherichia coli TITLE gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase J. Bacteriol. 172 (9), 5497-5500 (1990) JOURNAL 90368616 MEDLINE Location/Qualifiers **FEATURES** 1811..3109 Source /organism="Escherichia coli" /db_xref="taxon:562" 1811.. 1876 sig peptide /gene="appA" 1811..3109 CDS /gene="appA" /standard_name="acid phosphatase/phytase" /transl table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285" translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP/

/translation="MKAILIPFLSLLIPHTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG

ALEINWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSINT PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"

mat peptide

1877 3106 /gene="appA"

/product="periplasmic phosphoanhydride phosphohydrolase"

mutation replace(1817.. 1819, "gcg changed to gcc")

/gene="appA"

/standard_name="A3 mutant"

/note="created by site directed mutagenesis"

/phenotype="silent mutation"

mutation replace (3092..3094, ccg changed to ccc)

/gene="appA"

/standard_name=" P428 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

mutation replace (3095..3097, gcg changed to gct)

/gene="appA"

/standard_name=" A429 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

Figure 19 (continued):

BASE COUNT 1257 a 814 C 843 g 1146 t

ORIGIN

1 GGATCCCCTT TGCTATGTAG TTTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA

61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA 121 CTCTPTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT 181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG 241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG 301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA 361 TATTTCACTA AACTAGGTTT ATCTATTTTG TTGCTTTCTC TAACATCTCT GCAATGAAGC 421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TITCTACATA 481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TITCCAAGAA TGGAAAAGAA ATGTTCTGAC 541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT 601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG 661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA 721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT 781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT 841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT 901 TAAGATAAAG GTAACTGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG 961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG 1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT 1081 GATACTAACA CAGGTAAATC CCACACGTGT TITGAGACTA CATTTGCTGG GATTTTATTG 1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA 1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT 1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA 1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT 1381 TACTGATAAT ACAAACATGT GIGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA 1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TIATATAACG ATTGTATTTA TCAATCAATT 1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAATT TACTCTAATC 1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC 1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT 1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT 1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTTT CAGGAGCTAA GGAAGCTAAA 1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTTATCT CTTCTGATTC CGTTAACCCC 1861 GCAATCTGCA TTCGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG 1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCCAGA 1981 CECATGECCA ACCTEGCCEG TAAAACTEGE TTEGCTEACA CCECEGETE GTEAGCTAAT 2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA 2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTAA 2161 AACAGGCGAA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA 2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT 2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT 2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC 2401 ARACTTGTGC CTTARACGTG AGAAACAGGA CGAAAGCTGT TCATTARCGC AGGCATTACC 2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC 2521 AATGCTGACG GAGATATTTC TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG 2581 AAGGATCACC GATTCACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTTTA 2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCCGTTAT TAGATTTGAT 2701 CAAGACAGCG TTGACGCCCC ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC 2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA 2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGTT 2881, TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA

Figure 19 (continued):

```
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
    3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
    3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAG GTATAAGGCA
    3121 GTTATTGGTG CCCTTAAACG CCTGGTGCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
    3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
    3241 ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAAATAT AAAATTTTTA AGTGTATAAT
    3301 GTGTTAAACT ACTGATTCTA ATTGTTTGTG TATTTTAGAT TCCAACCTAT GGAACTGATG
    3361 AATGGGAGCA GTGGTGGAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
    3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA
    3481 GAAAGGTAGA AGACCCCAAG GACTTTCCTT CAGAATTGCT AAGTTTTTTG AGTCATGCTG
    3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC
    3601 TGCTATACAA GAAAATTATG GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
    3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
    3721 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT
    3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
    3841 TTACTTGCTT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
    3901 ATTGTTGTTG TTAACTTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC
    3961 ACAAATTTCA CAAATAAAGC ATTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
    4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC
//
```

Figure 20: Nucleic acid sequence of the known segment of the R15/appa plasmid (including the vector sequences of pBLCAT3 (SEQ ID NO:4).

LOCUS DNA SYN R15/appa 6116 bp 15-APR-2000 DEFINITION R15/appa transgene with vector ACCESSION R15/appa REFERENCE 1 (bases 1 to 6116) SOURCE synthetic construct. ORGANISM synthetic construct artificial sequence. **KEYWORDS** salivary proline-rich protein, acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence; AUTHORS Golovan, S., Forsberg, C.W., Phillips, J. **JOURNAL** Unpublished. DEFINITION Rat salivary proline-rich protein (RP15) gene. ACCESSION M64793 M36414 VERSION M64793.1 GI:206711 SOURCE Rat (Sprague-Dawley) liver DNA. ORGANISM Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. REFERENCE 1 (bases 1 to 1748) Lin, H.H. and Ann, D.K. AUTHORS TITLE Molecular characterization of rat multigene family encoding proline-rich proteins JOURNAL. Genomics 10, 102-113 (1991) MEDLINE 91257817 **FEATURES** Location/Qualifiers 1..1748 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db xref="taxon:10116" /tissue type="liver" /tissue lib="cosmid genomic library" misc_feature 1802-1810 /function=" consensus sequence for initiation in higher eukaryotes *

FEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375

VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1811.3109)

AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

TITLE The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase

and glucose-1-phosphatase

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

Figure 20 (continued):

```
MEDLINE 90368616
```

FEATURES

Location/Qualifiers

Source

1811..3109

Source

/organism="Escherichia coli"

/db_xref="taxon:562"

sig_peptide

/gene="appA"

CDS

1811..3109

1811.. 1876

/gene="appA"

/standard_name="acid phosphatase/phytase"

/transl_table=11

/product="periplasmic phosphoanhydride phosphohydrolase"

/protein_id="AAA72086.1" /db_xref="GI:145285"

/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP

TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS

GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA

NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS

ELKVSADNVSLTGAVSLASMLTEI FLLQQAQGMPEPGWGRI TDSHQWNTLLSLHNAQF

YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG

ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT

 ${\tt PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"}$

mat_peptide

1877 3106

/gene="appA"

/product="periplasmic phosphoanhydride phosphohydrolase"

mutation

replace(1817.. 1819, "gcg changed to gcc")

/gene="appA"

/standard_name="A3 mutant"

/note="created by site directed mutagenesis"

/phenotype="silent mutation"

mutation

replace(3092..3094, ccg changed to ccc)

/gene="appA"

/standard_name=" P428 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

mutation

replace (3095..3097, gcg changed to gct")

/gene="appA"

/standard_name=" A429 mutant"

/note="created by site directed mutagenesis"

/phenotype=" silent mutation "

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6116)

ACCESSION X64409

VERSION

X64409.1 GI:58163

SOURCE

synthetic construct.

ORGANISM synthetic construct

artificial sequence.

1 (bases 3109 to 6116)

REFERENCE AUTHORS

Luckow, B.H.R.

TITLE

Direct Submission

JOURNAL

Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res

Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG

Figure 20 (continued):

```
REFERENCE
                 2 (bases 3109 to 6116)
      AUTHORS
                Luckow, B. and Schutz, G.
      TITLE
                CAT constructions with multiple unique restriction sites
for
                 the functional analysis of eukaryotic promoters and
regulatory
                 elements
      JOURNAL
                Nucleic Acids Res. 15 (13), 5490 (1987)
      MEDLINE
                 87260024
     COMMENT
                 Promoterless CAT vector for transient transfection
experiments
                  with eukaryotic cells. Allows the analysis of foreign
                  promoters and enhancers.
     FEATURES
                          Location/Qualifiers
                          3109 to 6116
          source
                          /organism="synthetic construct"
                          /db xref="taxon:32630"
         polyA signal
                          3262..3457
                          /note="SV40 signals"
                          complement (4654..5514)
          CDS
                          /codon_start=1
                          /transl table=11
                          /gene="Amp"
                          /product="beta-lactamase"
                          /protein_id="CAA45753.1"
                          /db_xref="GI:58165"
BASE COUNT
               1724 a
                        1386 c
                                 1407 g
                                          1599 t
ORIGIN
        1 GGATCCCCTT TGCTATGTAG TTTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
       61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
      121 CTCTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
      181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
      241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
      301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
      361 TATITICACTA AACTAGGTTT ATCTATITTG TIGCTITCTC TAACATCTCT GCAATGAAGC
      421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
      481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
      541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
      601 TGGGAAGAAA CCATTTGGTG AACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
      661 TAAAACATAT GTTTGACCAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
      721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
      781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
      841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
      901 TAAGATAAAG GTAACTGTAT ACATTTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
      961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
      1021 TCCATGGACC TITGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCTT
      1081 GATACTAACA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTTGCTGG GATTTTATTG
     1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTCGGTTAC TTCAAGAGAA
     1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
     1261 GTARAGGAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
      1321 GTGTTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTTCAA TGGAAATAAT
     1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAATCCTA TCCAAATGCA CAGTGATACA
     1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
      1501 GTATGTATCA ATATATGGGC TATTTTCTTA CACATGATTT TATTCAAATT TACTCTAATC
      1561 ATTGTTGAAC CATTTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
      1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT
```

Figure 20 (continued):

1681	AGTATGAATT	CAAAGGTGCC	ATTCTTCTGC	TTCTAGTTAT	AAAGGCAGTG	CTTGCTTCTT
1741	CCAGCACAGA	TCTGGATCTC	GAGGAGCTTG	GCGAGATTTT	CAGGAGCTAA	GGAAGCTAAA
1801	AGCCGCCACC	ATGAAAGCCA	TCTTAATCCC	ATTTTTATCT	CTTCTGATTC	CGTTAACCCC
1861	GCAATCTGCA	TTCGCTCAGA	GTGAGCCGGA	GCTGAAGCTG	GAAAGTGTGG	TGATTGTCAG
1921	TCGTCATGGT	GTGCGTGCTC	CAACCAAGGC	CACGCAACTG	ATGCAGGATG	TCACCCCAGA
			TAAAACTGGG			
			AACGCCAGCG			•
			AGGTCGCGAT			
			CCGGGCTGGC			
			ATCCGTTATT			
			ACGCGATCCT			
			CGTTTCGCGA			
			AGAAACAGGA			
			CCGACAATGT			
			TCCTGCAACA			
		•	AGTGGAACAC			
			AGGTTGCCCG			
			ATCCACCGCA			
			GACACGATAC			
			GTCAGCCGGA			
			GCGATAACAG			
			ATAAAACGCC			
			GTGAAGAGCG			
			AAGCACGCAT			
			CCTGGTGCTA			
						TGAAACATAA
			TTAACTTGTT			•
			CAAATAAAGC			•
			CTTATCATGT			
			TGTTTCCTGT			
			TAAAGTGTAA			
			CACTGCCCGC			
			GCGCGGGGAG			
3721	TTCCTCGCTC	ACTGACTCGC	TECECTCEGT	CGTTCGGCTG	CGGCGAGCGG	TATCAGCTCA
			TATCCACAGA			
3841	AGCAAAAGGC	CAGCAAAAGG	CCAGGAACCG	TAAAAAGGCC	GCGTTGCTGG	CGTTTTTCCA
			AGCATCACAA			
3961	CCCGACAGGA	CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG	TGCGCTCTCC
4021	TGTTCCGACC	CTGCCGCTTA	CCGGATACCT	GTCCGCCTTT	CTCCCTTCGG	GAAGCGTGGC
4081	GCTTTCTCAA	TGCTCACGCT	GTAGGTATCT	CAGTTCGGTG	TAGGTCGTTC	GCTCCAAGCT
			CCGTTCAGCC			
			GACACGACTT			
4261	GATTAGCAGA	GCGAGGTATG	TAGGCGGTGC	TACAGAGTTC	TTGAAGTGGT	GGCCTAACTA
4321	CGGCTACACT	AGAAGGACAG	TATTTGGTAT	CTGCGCTCTG	CTGAAGCCAG	TTACCTTCGG
4381	AAAAAGAGTT	GGTAGCTCTT	GATCCGGCAA	ACAAACCACC	GCTGGTAGCG	GTGGTTTTTT
4441	TGTTTGCAAG	CAGCAGATTA	CGCGCAGAAA	AAAAGGATCT	CAAGAAGATC	CTTTGATCTT
4503	TTCTACGGG	TCTGACGCTC	AGTGGAACGA	AAACTCACGT	TAAGGGATTT	TGGTCATGAG
4561	ATTATCAAAA	AGGATCTTCA	CCTAGATCCT	TTTAAATTAA	AAATGAAGTT	TTAAATCAAT
			CTTGGTCTGA			
						TCGTGTAGAT
			TACCATCIGG			
						CCGAGCGCAG
4861	AAGTGGTCCT	GCAACTTTAT	CCCCCTCCAT	CCAGTCTATT	AATTGTTGCC	GGGAAGCTAG
						CAGGCATCGT
						GATCAAGGCG
						CTCCGATCGT
5101	TGTCAGAAGT	AAGTTGGCCG	CAGTGTTATC	ACTCATGGTT	ATGGCAGCAC	TGCATAATTC
			40160			

Figure 20 (continued):

```
5161 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
    5221 ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
    5281 TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG
    5341 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
    5401 CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
    5461 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
    5521 CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
    5581 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
    5641 ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC
    5701 GAGGCCCTTT CGTCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT
    5761 CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG
    5821 CGCGTCAGCG GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT
    5881 TGTACTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAATA
    5941 CCGCATCAGG CGCCATTCGC CATTCAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG
    6001 GGCCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGGGATGT GCTGCAAGGC GATTAAGTTG
    6061 GGTAACGCCA GGGTTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGC
//
```

Figure 21: Nucleic acid sequence of the known segment of the R15/appa transgene used for the generation of transgenic mice (SEQ ID NO:5).

```
LOCUS
            R15/appa
                        3470 bp
                                   DNA
                                                   SYN
                                                               15-APR-2000
DEFINITION
            R15/appa transgene with vector sequences removed.
ACCESSION
            R15/appa
REFERENCE
            1
               (bases 1 to 3470)
SOURCE
            synthetic construct.
       ORGANISM synthetic construct
                 artificial sequence.
KEYWORDS
            salivary proline-rich protein, acid glucose-1-phosphatase; appA
            gene; periplasmic phosphoanhydride phosphohydrolase; artificial
AUTHORS
            Golovan, S., Forsberg, C.W., Phillips, J.
  JOURNAL
            Unpublished.
     DEFINITION Rat salivary proline-rich protein (RP15) gene.
     ACCESSION
                 M64793 M36414
     VERSION
                 M64793.1 GI:206711
     SOURCE
                 Rat (Sprague-Dawley) liver DNA.
       ORGANISM
                 Rattus norvegicus
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
                 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
     REFERENCE
                 1 (bases 1 to 1748)
       AUTHORS
                 Lin, H.H. and Ann, D.K.
       TITLE
                 Molecular characterization of rat multigene family
encoding
                 proline-rich proteins
       JOURNAL
                 Genomics 10, 102-113 (1991)
       MEDLINE
                 91257817
     FEATURES
                           Location/Qualifiers
          source
                           1...1748
                           /organism="Rattus norvegicus"
                           /strain="Sprague-Dawley"
                           /db xref="taxon:10116"
                           /tissue_type="liver"
                           /tissue_lib="cosmid genomic library"
           misc feature
                                  1802-1810
                           /function=" consensus sequence for initiation in
                                       higher eukaryotes *
```

FEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375

VERSION M58708.1 GI:145283 SOURCE Escherichia coli DNA. ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1811..3109)

AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

TITLE The complete nucleotide sequence of the Escherichia coli gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase

44/58

```
J. Bacteriol. 172 (9), 5497-5500 (1990)
      JOURNAL
      MEDLINE
                90368616
FEATURES
                          Location/Qualifiers
                                 1811..3109
      Source
                          /organism="Escherichia coli"
                          /db_xref="taxon:562"
                          1811.. 1876
      sig_peptide
                          /gene="appA"
                           1811..3109
      CDS
                           /gene="appA"
                          /standard_name="acid phosphatase/phytase"
                          /transl_table=11
                          /product="periplasmic phosphoanhydride phosphohydrolase"
                           /protein_id="AAA72086.1"
                           /db xref="GI:145285"
/translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
GQVAI IADVDERTRKTGEAFAAGLAFDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
nvtdailsraggsiadftghrqtafrelervinfpqsnlclkrekqdescsltqalps
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YILQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
                           PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL*
                                 3106
                           1877
       mat_peptide
                           /gene="appA"
                           /product="periplasmic phosphoanhydride phosphohydrolase"
                     replace(1817.. 1819, "gcg changed to gcc")
       mutation
                     /gene="appA"
                     /standard name="A3 mutant"
                     /note="created by site directed mutagenesis"
                     /phenotype="silent mutation"
                     replace(3092..3094, " ccg changed to ccc")
       mutation
                     /gene="appA"
                     /standard_name=" P428 mutant"
                     /note="created by site directed mutagenesis"
                     /phenotype=" silent mutation "
                     replace (3095..3097, * gcg changed to gct*)
        mutation
                     /gene="appA"
                     /standard name=" A429 mutant"
                     /note="created by site directed mutagenesis"
                     /phenotype=" silent mutation "
                              3262..3457
            polyA signal
                              /note="SV40 signals"
                                                 949 t
  BASE COUNT
                  1065 a
                             721 C
                                       735 g
  ORIGIN
          1 GGATCCCCTT TGCTATGTAG TTTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
         61 GAGAGTCCTG TTTGGTTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
        121 CTCTTTGTTT CTAGCATAAC CAAAAGATTT AGTGAATTGA AAACAATGTT CCCTTAGAGT
        181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTIGTAAG TATCTCATAG
        241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTTAA CATGATTTTC ATTAATCAGG
```

//

					ATAGTGGTCC	
					TAACATCTCT	
					CACAAAAAAT	-
					TGGAAAAGAA	
					AGGCACACAA	
					TAGACAAACA	•
					AATGTGAATA	
					AAATATAAA	
					ATCTTGAGAG	
					CCTGAGCTAT	
					ACAAGAAAGC	
961	TTCAGCTCTA	TAATTCTTGC	CTTAAACAAC	TTAAATAGAA	TGATTTAAAA	TATGGAGCTG
					CAAGGAATTA	
1081	GATACTAACA	CAGGTAAATC	CCACACGTGT	TITGAGACTA	CATTIGCTGG	GATTTTATTG
1141	ATGTAATAGG	TCACATGTTT	TTCGGGCCAA	TGTTGCTGTT	ATTCGGTTAC	TTCAAGAGAA
1201	TAGTGGCAAC	TGATGCTATG	TATTCTAGGG	GTTTGAAGTG	ATGTTTCATG	ATIGAAATIT
1261	GTAAAAGAAT	AACATCATCA	TTCTTAACAA	TAGAACATAT	AAAGTCACAC	AGAAGTGACA
1321	GIGTTTAAGC	TGTACTATTG	ATCAAAGAAA	TTTATTACCT	TCAGTTTCAA	TGGAAATAAT
					TCCAAATGCA	
1441	CAGAAAATAT	TAGCAAGTAG	AATGCAATAT	TTATATAACG	ATTGTATTTA	TCAATCAATT
					TATTCAAATT	
					CTTACCTCAT	-
					TCTGTTAAGT	
					AAAGGCAGTG	
					CAGGAGCTAA	
					CTTCTGATTC	
	·				GAAAGTGTGG	- · · · · ·
					ATGCAGGATG	
					CCGCGCGGTG	
					GACGGATTGC	
					GTCGACGAGC	
					GCAATAACCG	
					AAAACTGGCG	
					GGAGGGTCAA	
					GTGCTTAATT	
					TCATTAACGC	
					GGTGCGGTAA	
					ATGCCGGAGC	
					TTGCATAACG	
					ACCCCGTTAT	
					TATGGTGTGA	
					AATCTCGGCG	
					CCAGGTGGTG	
					CAGGTTTCGC	
					AATACGCCGC	
					GGCATGTGTT	
					AGTTTGTAAG	
					AGTGATAATA	
					CIGIGGIGIG	
					CCCCTGAACC	
					TATAATGGTT	
					CTGCATTCTA	
3421	GTCCAAACTC	ATCAATGTAT	CTTATCATGT	CTGGATCGAT	CCCCGGGTAC	

Figure 22: Nucleic acid sequence of the SV40/APPA+intron plasmid (SEQ ID NO:6).

CIRCULAR SYN 14-APR-2000 DNA 5421 bp LOCUS SV40/APPA DEFINITION Ligation of SV40 promoter/enhancer into CAT/APPA+intron ACCESSION SV40/APPA REFERENCE 1 (bases 1 to 5421) synthetic construct. ORGANISM synthetic construct artificial sequence. SV40 promoter/enhancer, acid glucose-1-phosphatase; appA gene; KEYWORDS periplasmic phosphoanhydride phosphohydrolase; artificial sequence; Golovan, S., Forsberg, C.W., Phillips, J. **AUTHORS** JOURNAL Unpublished. DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene. M58708 L03370 L03371 L03372 L03373 L03374 L03375 ACCESSION M58708.1 GI:145283 VERSION Escherichia coli DNA. SOURCE ORGANISM Escherichia coli Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia. (bases 40 1337) REFERENCE 1 AUTHORS Dassa, J., Marck, C. and Boquet, P.L. The complete nucleotide sequence of the Escherichia coli gene appA TITLE reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990) MEDLINE 90368616 Location/Qualifiers FEATURES 40 1337 Source /organism="Escherichia coli" /db xref="taxon:562" 40.. 105 sig_peptide /gene="appA" . 40 1337 CDS /gene="appA" /standard_name="acid phosphatase/phytase" /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAA72086.1" /db_xref="GI:145285" /translation="MKAILIPFLSILIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS GQVALIADVDERTRKTGEAFAAGLAPDCALTVHTQADTSSPDPLFNPLKTGVCQLDNA NVTDAILSRAGGS LADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPS elkvsadnvsltgavslasmlteifllqqaqgmpepgwgritdshqwntllslµnaqf YLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLF1AGHDTNLANLGG ALELNWILPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL* 1334 106 mat_peptide /gene="appA"

47/58

Figure 22 (continued):

```
/product="periplasmic phosphoanhydride phosphohydrolase"
      mutation
                   replace(46.. 48, "gcg changed to gcc")
                   /gene="appA"
                   /standard_name="A3 mutant"
                   /note="created by site directed mutagenesis"
                   /phenotype="silent mutation"
      mutation
                   replace(1320..1322, " ccg changed to ccc")
                   /gene="appA"
                   /standard_name=" P428 mutant"
                   /note="created by site directed mutagenesis"
                   /phenotype=" silent mutation *
      mutation
                   replace(1323..1325, gcg changed to gct*)
                   /gene="appA"
                   /standard name=" A429 mutant"
                   /note="created by site directed mutagenesis"
                   /phenotype=" silent mutation "
DEFINITION Plasmid pBLCAT3 (bases 2200 to 4924)
     ACCESSION
                X64409
     VERSION
                 X64409.1 GI:58163
     SOURCE
                 synthetic construct.
       ORGANISM synthetic construct
                 artificial sequence.
     REFERENCE
                 1 (bases 2200 to 4924)
       AUTHORS
                 Luckow, B.H.R.
       TITLE
                 Direct Submission
       JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
                  Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
     REFERENCE
                 2 (bases 2200 to 4924)
       AUTHORS
                 Luckow, B. and Schutz, G.
       TITLE
                  CAT constructions with multiple unique restriction sites
for
                  the functional analysis of eukaryotic promoters and
regulatory
                  elements
       JOURNAL
                  Nucleic Acids Res. 15 (13), 5490 (1987)
       MEDLINE
                  87260024
     COMMENT
                  Promoterless CAT vector for transient transfection
experiments
                   with eukaryotic cells. Allows the analysis of foreign
                   promoters and enhancers.
      FRATURES
                           Location/Qualifiers
           source
                           2200 to 4924
                            /organism="synthetic construct"
                            /db_xref="taxon:32630"
           SV40 t intron
                           1380..1993
                            /note="SV40 signals"
           polyA_signal
                           1990..2230
                            /note="SV40 signals"
           CDS
                            complement (3471..4317)
                            /codon start=1
                            /transl_table=11
                            /gene="Amp"
                            /product="beta-lactamase"
                            /protein_id="CAA45753.1"
                            /db_xref="GI:58165"
```

48/58

PCT/CA00/00430

Figure 22 (continued):

SV40 promoter/enhancer 5023..5402 /note="SV40 signals"

1331 g 1321 C BASE COUNT 1413 a 1 CGAGATTITC AGGAGCTAAG GAAGCTAAAA GCCGCCACCA TGAAAGCCAT CTTAATCCCA ORIGIN 61 TTTTTATCTC TTCTGATTCC GTTAACCCCG CAATCTGCAT TCGCTCAGAG TGAGCCGGAG 121 CTGAAGCTGG AAAGTGTGGT GATTGTCAGT CGTCATGGTG TGCGTGCTCC AACCAAGGCC 181 ACGCAACTGA TGCAGGATGT CACCCCAGAC GCATGGCCAA CCTGGCCGGT AAAACTGGGT 241 TGGCTGACAC CGCGNGGTGG TGAGCTAATC GCCTATCTCG GACATTACCA ACGCCAGCGT 301 CTGGTAGCCG ACGGATTGCT GGCGAAAAAG GGCTGCCCGC AGTCTGGTCA GGTCGCGATT 361 ATTGCTGATG TCGACGAGCG TACCCGTAAA ACAGGCGAAG CCTTCGCCGC CGGGCTGGCA 421 CCTGACTGTG CAATAACCGT ACATACCCAG GCAGATACGT CCAGTCCCGA TCCGTTATTT 481 AATCCTCTAA AAACTGGCGT TTGCCAACTG GATAACGCGA ACGTGACTGA CGCGATCCTC 541 AGCAGGGCAG GAGGGTCAAT TGCTGACTTT ACCGGGCATC GGCAAACGGC GTTTCGCGAA 601 CTGGAACGGG TGCTTAATTT TCCGCAATCA AACTTGTGCC TTAAACGTGA GAAACAGGAC 661 GAAAGCTGTT CATTAACGCA GGCATTACCA TCGGAACTCA AGGTGAGCGC CGACAATGTC 721 TCATTAACCG GTGCGGTAAG CCTCGCATCA ATGCTGACGG AGATATTTCT CCTGCAACAA 781 GCACAGGGAA TGCCGGAGCC GGGGTGGGGA AGGATCACCG ATTCACACCA GTGGAACACC 841 TTGCTAAGTT TGCATAACGC GCAATTTTAT TTGCTACAAC GCACGCCAGA GGTTGCCCGC 901 AGCCGCGCCA CCCCGTTATT AGATTTGATC AAGACAGCGT TGACGCCCCA CCACCGCAAA 961 AACAGGCGTA TGGTGTGACA TTACCCACTT CAGTGCTGTT TATCGCCGGA CACGATACTA 1021 ATCTGGCAAA TCTCGGCGGC GCACTGGAGC TCAACTGGAC GCTTCCCGGT CAGCCGGATA 1081 ACACGCCGCC AGGTGGTGAA CTGGTGTTTG AACGCTGGCG TCGGCTAAGC GATAACAGCC 1141 AGTGGATTCA GGTTTCGCTG GTCTTCCAGA CTTTACAGCA GATGCGTGAT AAAACGCCGC 1201 TGTCATTAAA TACGCCGCCC GGAGAGGTGA AACTGACCCT GGCAGGATGT GAAGAGCGAA 1261 ATGCGCAGGG CATGTGTTCG TTGGCAGGTT TTACGCAAAT CGTGAATGAA GCACGCATAC 1321 CCGCTTGCAG TTTGTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGTGCT ACGCCTGAAT 1381 AAGTGATAAT AAGCGGATGA ATGGCAGAAA TTCGCCGGAT CTTTGTGAAG GAACCTTACT 1441 TCTGTGGTGT GACATAATTG GACAAACTAC CTACAGAGAT TTAAAGCTCT AAGGTAAATA 1501 TAAAATTITT AAGTGTATAA TGTGTTAAAC TACTGATTCT AATTGTTTGT GTATTTTAGA 1561 TTCCAACCTA TGGAACTGAT GAATGGGAGC AGTGGTGGAA TGCCTTTAAT GAGGAAAACC 1621 TGTTTTGCTC AGAAGAAATG CCATCTAGTG ATGATGAGGC TACTGCTGAC TCTCAACATT 1681 CTACTCCTCC AAAAAAGAAG AGAAAGGTAG AAGACCCCAA GGACTTTCCT TCAGAATTGC 1741 TAAGTTTTTT GAGTCATGCT GTGTTTAGTA ATAGAACTCT TGCTTGCTTT GCTATTTACA 1801 CCACAAAGGA AAAAGCTGCA CTGCTATACA AGAAAATTAT GGAAAAATAT TCTGTAACCT 1861 TTATAAGTAG GCATAACAGT TATAATCATA ACATACTGTT TTTTCTTACT CCACACAGGC 1921 ATAGAGTGTC TGCTATTAAT AACTATGCTC AAAAATTGTG TACCTTTAGC TTTTTAATTT 1981 GTAAAGGGGT TAATAAGGAA TATTTGATGT ATAGTGCCTT GACTAGAGAT CATAATCAGC 2041 CATACCACAT TTGTAGAGGT TTTACTTGCT TTAAAAAACC TCCCACACCT CCCCCTGAAC 2101 CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT TTATTGCAGC TTATAATGGT 2161 TACAAATAAA GCAATAGCAT CACAAATTIC ACAAATAAAG CATTTTTTIC ACTGCATICT 2221 AGTTGTGGTT TGTCCAAACT CATCAATGTA TCTTATCATG TCTGGATCGA TCCCCGGGTA 2281 COGAGCTOGA ATTOGTAATO ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC 2341 ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA 2401 GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG 2461 TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGA GAGGCGGTTT GCGTATTGGG 2521 CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG 2581 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA 2641 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG 2701 GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG 2761 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC 2821 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG 2881 GGAAGCGTGG CGCTTTCTCA ATGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT 2941 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG CGCCTTATCC 3001 GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC 3061 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG

Figure 22 (continued):

```
3121 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
3181 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
3241 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT
3301 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCACG TTAAGGGATT
3361 TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT
3421 TITARATCAR TOTARAGTAT ATRIGAGTAR ACTIGGTOTG ACAGTTACCA ATGCTTRATC
3481 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGTTGC CTGACTCCCC
3541 GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA
3601 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
3661 GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC
3721 CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT
3781 ACAGGCATCG TGGTGTCACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC CGGTTCCCAA
3841 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT
3901 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA
3961 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC
4021 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA
4081 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTITAAAAG TGCTCATCAT TGGAAAACGT
4141 TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC
4201 ACTCGTGCAC CCAACTGATC TTCAGCATCT TITACTTTCA CCAGCGTTTC TGGGTGAGCA
4261 AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
4321 CTCATACTCT TCCITTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC
4381 GGATACATAT TIGAATGTAT TIAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC
4441 CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC CTATAAAAAT
4501 AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA
4561 CACATGCAGC TCCCGGAGAC GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA
 4621 GCCCGTCAGG GCGCGTCAGC GGGTGTTGGC GGGTGTCGGG GCTGGCTTAA CTATGCGGCA
 4681 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA
 4741 AGGAGAAAAT ACCGCATCAG GCGCCATTCG CCATTCAGGC TGCGCAACTG TTGGGAAGGG
 4801 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGGATG TGCTGCAAGG
 4861 CGATTAAGTT GGGTAACGCC AGGGTTTTCC CAGTCACGAC GTTGTAAAAC GACGGCCAGT
 4921 GCCAAGCTTT ACACTTTATG CTTCCGGCTC GTATGTTGTG TGGAATTGTG AGCGGATAAC
 4981 AATTTCACAC AGGAAACAGC TATGACCATG ATTACGAATT CGGCGCAGCA CCATGGCCTG
 5041 AAATAACCTC TGAAAGAGGA ACTTGGTTAG GTACCTTCTG AGGCGGAAAG AACCAGCTGT
 5101 GGAATGTGTG TCAGTTAGGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
 5161 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG
 5221 GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC
 5281 CGCCCATCCC GCCCCTAACT CCGCCCAGTT CCGCCCCATTC TCCGCCCCAT GGCTGACTAA
 5341 TITTTTTAT TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT
 5401 GAGGAGGCTC GAGGAGCTTG G
```

//

Figure 23. The nucleic acid sequence of the Lama2/APPA transgene used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7)

LOCUS transgene 17732 bp DNA SYN 14-APR-2000 DEFINITION Lama-appA cut XhoI..20623 to NotI..17732 ACCESSION transgene KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA periplasmic phosphoanhydride phosphohydrolase; artificial sequence; cloning vector REFERENCE 1 (bases 1 to 17732) **AUTHORS** Golovan, S., Forsberg, C.W., Phillips, J. JOURNAL Unpublished. FEATURES DEFINITION M. musculus Psp gene for parotid secretory protein. ACCESSION X68699 VERSION X68699.1 GI:53809 SOURCE house mouse. ORGANISM Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. REFERENCE 1 (bases 3777 to 5332;) AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P. TITLE Novel salivary gland specific binding elements located in the PSP proximal enhancer core Nucleic Acids Res. 26 (11), 2761-2770 (1998) JOURNAL MEDLINE 98256451 REFERENCE 2 (bases 7147 to 12653; 13952 to 17731) AUTHORS Mikkelsen, T.R. TITLE Direct Submission JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular Biology, University of Aarhus, CF Mollers Alle 130, 8000 Aarhus, DENMARK REFERENCE 3 (bases 7147 to 12653; 13952 to 17731) Laursen J, Hjorth JP TITLE A cassette for high-level expression in the mouse salivary glands. JOURNAL. Gene 1997 Oct 1;198(1-2):367-72 MEDLINE 9370303 **FEATURES** Location/Qualifiers source 1.to 12653; 13952 to 17731 /organism="Mus musculus" /strain="C3H/As" /db_xref="taxon:10090" /chromosome="2" /map="Estimate: 69 cM from centromere" /clone="Lambda YP1, Lambda YP3, Lambda YP7" /clone_lib="Lambda-PHAGE (Lambda L47.1)" /germline /note="Allele: b" misc feature 3777-5332 /gene="PSP" /function="salivary gland specific positive acting regulatory region" enhancer 7147..8724

51/58

Figure 23 (continued):

```
/evidence=experimental
     exon
                  11778..11824
                  /qene="Psp"
                  /note="exon a"
                  /number=1
                  /evidence=experimental
     aoxe
                  12626.. 14190
                  /gene="Psp"
                  /note="exon b fused with exons h and i"
     misc feature
                        12644-12652
                  /function=" consensus sequence for initiation in higher
                  eukaryotes · "
     misc feature
                        13952-13965
                  /function=" M13mp18 polylinker"
DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
gene,
     ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
      VERSION
                  M58708.1 GI:145283
      SOURCE
                       Escherichia coli DNA.
      ORGANISM
                 Escherichia coli
           Bacteria; Proteobacteria; gamma subdivision;
      Enterobacteriaceae;
            Escherichia.
REFERENCE
            1 (bases 12653..13951)
               Dassa, J., Marck, C. and Boquet, P.L.
      AUTHORS
                The complete nucleotide sequence of the Escherichia coli
      TITLE
                  gene appA reveals significant homology between pH 2.5
                  acid phosphatase and glucose-1-phosphatase
                J. Bacteriol. 172 (9), 5497-5500 (1990)
      JOURNAL.
      MEDLINE
                90368616
FEATURES
                        Location/Qualifiers
      Source
                              12653..13951
                         /organism="Escherichia coli"
                        /db_xref="taxon:562"
                        12653..12718
      sig_peptide
      /gene="appA"
      CDS
                        12653
                                     13951
                         /gene="appA"
                         /standard_name="acid phosphatase/phytase"
                         /transl_table=11
                         /product="periplasmic phosphoanhydride
                        phosphohydrolase*
                         /protein id="AAA72086.1"
                         /db xref="GI:145285"
 /translation="MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAP
TKATQLMQDVTPDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQS
```

52/58

GOVAI IADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA

NVTDAILSRAGGSIADFTGHROTAFRELERVLNFPOSNLCLKREKODESCSLTOALPS

RECTIFIED SHEET (RULE 91)

Figure 23 (continued):

```
ELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQF
YLLORTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGG
ALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNT
PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL"
     mat_peptide
                      12719 13948
                       /gene="appA"
                       /product="periplasmic .phosphoanhydride
                            phosphohydrolase"
                 replace(12659.. 12661, "gcg changed to gcc")
     mutation
                 /gene="appA"
                 /standard name="A3 mutant"
                 /note="created by site directed mutagenesis"
                 /citation=[3]
                 /phenotype="silent mutation"
     mutation
                 replace (13934..13936, " ccg changed to ccc")
                 /gene="appA"
                 /standard name=" P428 mutant"
                 /note="created by site directed mutagenesis"
                 /citation=[3]
                 /phenotype=" silent mutation "
                 replace(13937..13939, * gcg changed to gct*)
      mutation
                 /gene="appA"
                 /standard name=" A429 mutant"
                 /note="created by site directed mutagenesis"
                 /citation={3}
                  /phenotype=" silent mutation "
BASE COUNT
               4719 a
                       4125 C
                                4168 g
                                        4719 t
ORIGIN
        1 TCGAGAGTAT CTTTGTCAGC TGTGCCTCCA ACAAAGGGGT ACTGTTGCCC ACATAGAAAG
       61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
      121 TGTTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
      181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
      241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
      301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
      361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
      421 AGGGATTGAG AGACCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACTTG
      481 ACAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGACCCG
      541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCAGTATT
      601 ATAAAAGGAC AGTATTACAG ATTTTGTTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
      661 TTTAAGTAGG GTAAAGTACT CTITAAAAAT GGGTCCTAGA TATTTTTTCC TTTAACTCAA
      721 GTCTCTTACT GTTTAAATGA TTTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
      781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
      841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
      961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
      1021 GCGGGGCGTG GTGGCATACA CTTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
      1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
      1201 ACCAAACCAA ACCAAACCAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
```

1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCACAGC CAGGCTGTGG AACCAGCCTG AGTGTCCATG ATAAATGAAT
53/58

1381	GGATAGGTAA	CTTTCAAGGT	AAATGGACTC	TGCTGTGTAC	ATGCCTCACA	TTCTGTTTAT
1441	TCATTTTTCT	TTATGAGGTG	TCCATTCAGG	AGTCACATGG	TAGTTCTATT	TTCAGTCTTC
1501	TGAAGATACT	ACACTGGTCC	CCACAGTTTA	CACTTTTATC	AGCAGTGAAT	AAGGGTTCCT
1561	CTATCCTTAC	CATCATTIGT	TGTAATTTTT	CTTGATGACC	CTCTTTCTGA	CAGGGATAGG
1621	ATGTAATATC	AGTGTGAGGA	AGTACAACTT	GTTTTCTAAG	TATTTATTGG	CCCCTTGCAT
1681	TTCTTCTTTT	GAAAACTGTC	GGTTCCTGAC	ATCTGCTCAG	GTATTCATTG	GAŢGTTGTTT
1741	CITTGGTGTT	TGAGTTCTTA	TGAATTCTAG	ATGTTAAATC	CCTGCCTGTG	GTTCTCTCCC
1801	ATTCTGTAGG	CTGCCTCCTC	ACCCTGGCAA	TTGTTGTCCT	TGTTTTGCAG	AAACTTTTGA
1861	CTTCATGGAA	TCTCATTTGT	CAGTTTTCCC	TCCTCTGCTA	TAGCCTGAGC	TAATGCACTG
1921	GTTTTTACAG	AGCCCTGGTC	TATGCCTTTA	TCCTCCTCTG	GCAGCTTCGG	AGTTTCATTT
1981	CTTACATTTA	GATCTTTGAT	CCACTTTGAA	CAAGTTTTGG	AGCAGGGTGA	GAGATACGAA
2041	TCTAGTTCCA	TTCTTCCATA	TGTGATCCTA	GTTTACATAG	CATCGTTGGT	TGAAGAGGTT
2101	TTATTTTATT	TTTAAATAAT	GTGTCATAAA	AAACGAGGTG	GTTGTAGCAG	TGTGGATTTG
2161	TTTCTTTGTC	CTITGATCIA	CAGGTCTTGT	TTTGTGTCAG	TCTCATGATG	TITTATIGCT
2221	ATGGCTCTGT	CATACAGTCT	GAGGTCAGGT	ATTGTGATAT	ACCTTCAGTA	TIGCTCCCTC
2281	AGACTCAGGT	TTGCTTTGGC	CAGGAGTCAT	CTTACTCAGT	GCTCTTAGAG	CTCCCCCAGC
2341	ATGTAGCTGC	TACTATTCTT	AGTTGATAAA	TCAGGAAACT	GGGGCTCAGA	GAGATTAACT
2401	GTCTTGAACT	ACTTCTGGGG	AGGTGAAACG	TGGAGACACT	AAACTGTGTT	TACCCTGTAC
2461	TGCTCCAGTA	GCTGTCGGGT	GCTGGGCTAC	AGCAAAGCAC	CTATACTATA	TATTACTCAG
2521	GAGGTGGAAA	AACTCAGCCT	CCCTTGGGGT	TCCCAAGCTC	CCAGGTGTCC	AGTCACTGCT
2581	GGAAACCTCA	TGGAGTCTGA	AAGGAAGGGT	TGAGGGTACA	TGGGGCAGCG	ATGAGGAGCC
2641	TREGGETTEGE	ATCTCCCAAA	CACCTGGATA	TCCAGATGCC	ACTGGGTCAG	GGGGAGTTGG
2701	GAACAGAGTT	GGGATGTCCA	TGGACCTGTG	ACAAGGCCAG	GGCCAGGGGG	AGGATAACTC
2761	AUTCHALD (AL	AATTTGCGAA	AGTCCTTAGC	TTAGCAGCAG	TTGTCTGGGA	GCACAGAGGG
2821	COUTTOTA	AGAGGCTCAG	GCAGTGCCGC	TCTGTAGGCG	AAGGTCTTCT	CCATGTTCCC
2881	CATGGTGGTT	CTTGATGAAA	GAGACAGTCC	TTGGCTCCAA	ACTGGTTTAT	TGATTGTTCA
2941	TTCTCCAAAA	TGGGTGCACA	CCACCTTCTC	AGGGTGGACC	AGAGATCAAA	TACCTTTTGC.
3001	AGGGAGGAAT	ATCTGGGAAG	GGACGCTTAC	TGGCTAAACC	CTCAGGGCCT	CTAGATACAT
3061	CATTAGCATG	GAGAACTCTG	TTCTGGGCTA	CATGACCACA	GGCCACATTT	CCACAAGCCA
3121	CATGTGGGAA	GTGTGGCACA	TGTTCTAGGC	CAGGAATCTG	GTAGGGAGCG	TGGAGCCACC
3181	TACCATCCCA	GGTGGGTGCC	TGGGTGCCAG	GGACCCTGAA	CCCGCTCAAC	CTTACCAAGT
3241	TTCCTGGCAG	GGTCCACTGT	CCTACACAGA	AGCTGGAGGA	GGTGTGAGGG	TIGTGTCTTT
3301	GTGGAATGTC	CCATGCTGCT	TGGGGCTCAG	TTTCTCCACC	TGTACCTCAT	TGGTTTGGGT
3361	ATAAAAAGTG	GGGATACTTT	ATTATTCTCT	GACTCGGTCC	TGAGGAAAAA	GCATCGTGGC
3421	AGTCCAGGAA	CCACACCCTG	AGGTTCCTGC	ACTGAAGGGA	CTCCCTAAGT	CTCTGGAGTC
3481	TCTCCCCTTC	ACAGAGCTGC	CAAAGTCTAG	GTTCTTTTGA	GGATAACAGA	GCCATGCTTG
3541	GTAAGCAGAC	AACAGCATTT	GTTTACTCAA	CCTTCTTTTG	TCAGCTCCCT	CTTCATAAAC
3601	AAGTTGAGAC	ACCATGCTGG	CTTGAGGAAG	ACTTCTAAAG	CCAGACAACT	GTGCAAGGAA
3661	GAAGAAGAAG	GGGCAAGTGG	AGTTAGCCTG	GATGTAGCCC	TCAAAGTCTC	CAGAGACCAG
372	CCATGAAGGC	TCAAGTGGAG	GGCAAGACCT	GCAGCAGCCA	AGCATCTGGC	AGGAGAGGAT
3783	CCTGGGAACC	CCTCTACCAT	GACACACATT	CITCCTGCAG	GTCACACTTA	ATAGGCCATT
3847	TCTTATTTG	ATCTATCATC	GTGTTCTGTG	CGAGATTAAT	GAGGTGTTAT	GCTGCGAACA
390	I GARAGITATA	TAAAAACAAG	TCCCCCCCC	TTGTCACTGC	TGCTAAGAAT	GTAGCAGAAA
396	1 TTGTCTCAAG	TGTCTCTCTA	ATCAGAAACA	ATAAAGGTCI	CCTTGGATTC	AAGCCCTCCA
402	1 GTTTCCTCCT	TCCTTGCTGA	GCCTTGGACA	CCCATACAAA	CCTCCTGGAT	GCTACAGCTC
40B	1 TGGGCAGAGA	CTCCAAGGTG	GGGAGAGACT	GATGGTACAA	AAGCAAAATA	CTIGTITGGG
414	1 GGTACACCC	CTCCTCTGCC	TGTGTGGTTC	CTGCAGTCAG	TCCTGCAGAC	AGGCCCTCAG
420	1 TGGGTCTTC	ATGGGCAACA	CGCAGAGGGA	GGCAATGGAT	GGGAATACCC	ACACCCTGGT
426	1 TAGTTTACCO	CGGCCATGCT	CTCTGCTCTT	CATCCCTCCI	CTGCCCTCTG	CCACGGCTTT
432	1 CTCTGCAGG	ATCATATCT	CATATTGGCC	CACAGGTGTT	CTCCTCACCC	TAGCTATGAT
438	1 GTTTACTTT	A GAGTGACCT	AGCAGGGCTG	GTGGGAATG	GTTCTAGAAG	GCTCACGGAG
444	1 ATGCTAGGG	AGAAACGTC	TCTAACTACT	GAGGTTACT	AGTICCIGGI	GGTTGTCTCT
450	1 6000000	TGTTAAAGT	ACCTTGAAGT	TAGTGCAGA	GAAATCAGAG	CCCAGTCACA
456	1 GAGTAAATA'	r GGTCCTGAAC	ATTTCCTTTC	AGTGCCCAG	ATCCATGACA	TTTCAAGAGC
467	1 CCalcaladaticata	A CCTTAAGTC	TTTGGGGTTC	TATCTTCTG	TTGATGTATG	TCTCTCTCTT
460	1 TATCALAGA	G TGAGATGGT	CATAAGAGG	TGCTCTAAA	GACAGAGAG	ATTIGCAATT
474	1 GTGGCATGT	G ACATCCTCAC	GCCTTGCTC	GGTGCCAGG	GGAACTGATG	CAGAAAAGAG
480	1 TADGAGGTC	A TTTCCTGGA	GCTGTCACT	TAGAGGAGAT	CTTACAGTGC	ATTCCCTCCT
-10V	- INDOMOUIC	. ILICIOM				

4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG
The same and the s
THE TAXABLE TO A COCCE CACCECTORY CHICKITATIC CALABOLITIC TAXABLE
THE PARTY OF THE P
THE PARTY OF THE PROPERTY OF A CONTROL OF THE PROPERTY OF THE
THE PARTY OF THE P
THE PARTY WAS CONCOUNTED BY CO
THE PROPERTY OF THE PROPERTY O
The state of the s
TO THE PROPERTY OF ACADEMY AND A CONTRACTOR OF THE PROPERTY OF
THE PERSON OF TAKEN AND A PROPERTY OF THE PERSON OF THE PE
THE PART OF THE PROPERTY AND
TO THE REPORT OF THE PROPERTY
THE COMMENT OF SUMMER PROPERTY OF THE PROPERTY
THE TANK AND COMPANY OF THE PARTY OF THE PAR
ACAMPATOR CAMPITALIST CALCALIST TO THE CONTROL OF THE CALCALIST TO THE CAL
The state of the s
The second secon
The second of th
THE REPORT OF THE PART OF THE
THE PROPERTY OF A STANDARD AND A STANDARD COLORS OF THE STANDARD COL
The second design of the second secon
The state of the s
The second of the confidence o
6901 TITCCCCTCA CACCUTCACT CHARTECTCA TGCTTCTCTC CTCCCTCAAA TACTTGTCAC 6961 TGTGGTCCCT CCACTTTCCT TTATCTCTCA TGCTTCTCTC CTCCCTCAAA TACTTGTCAC
6961 TGTGGTCCCT CCACTITICCT TIATCTGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC 7021 CCACTATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGGTTC
7021 CCACTATACT TCAGGGGCCA GCTCTACCA AGAACTCTCT CCAGGTCCAA TTTCAGGTTC 7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAACTCTCT CCAGGTCCAA TTTCAGGTTC
7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA ACATGCCACT AACAAGAAGA TGCAAATTCC 7141 CTTCGAAAGC TTTCAGCAAA TGCTCAGGGA ACATGCCACT AACAAGAAGA TGCAAATTCC 7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCATCTTC CAGGCCAAGG TCAGAGGGGC 7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCATCTTC CAGGCCAAGG TCAGAGGGGC
7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCATGTTTT GTTTTTAAGG TTTATTTATT 7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTTT GTTTTTAAGG TTTATTTATT
7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTGCAT CATGTGTGTG CAGGTTCCTG 7321 TTAGGTGTTA GTGTCTTTGC TTGCATGACC TTATGTGCAT CATGTGTGTG CAGGTTCCTG
7381 ATGACAGTAG AGGAGGGCTT TGACTCCTG GAAGTGAATGC TCTTCACCAC TGAGCCATCT 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAC TGAGCCATCT 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAC TGAGCCATCT 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAC TGAGCCATCT 7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG TATAATTGTG TGTTGTGTGT AACATTGAGG
7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AACATTGTG TGTGTGTGTC AACATTGAGG 7501 CTCTAGGCCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTGTGTCT AACATTGAGG
7501 CTCTAGGCCC AAGAGACATT GCTTTATGGA TALAGGTTGTA CAGTTTCACT AATTGCTACT 7561 AAAGGGAAAT AAAAAAAAAA CTTCAGCCGC TAAGGTTGTA CAGTTTCACT AATTGCTACT
· and managed which could be a facility of the second recommendation
8221 CAACAGCCAA GTATTTTCCA TIAGAGGAGA CITCUTOTA CITGACAGCCT AACCTCTCCC 8281 AAGGTGACTT GGGGCAGTCA GTACAGACTT GGGATGACCT CTGACAGCCT AACCTCTCCC
8281 AAGGTGACTT GGGGGGGGTCA GTACAGACTT GGGATCAG
<u> </u>

				•		
8341	CAACAAGGGC	CCTCTATGTT	TGCTATGTAA	TGTAATGTCA	GACATTGTCA (GGAGTGTCCG
8401	CAGCACAGCC	TGCCCAGTGT	GAGGGCTCTC	ATAGGTTTCC	CACTGTCTTA '	TCTACACAGG
8461	GATAACGAGG	AGGTAAGCTG	CAGTTCCCAG	TCTCACTTCA	CAGAGGAAGA (GATAACCCCA
8521	TCCCAGGTCA	TGTAGCCAGC	agtggaaaga	ATGAGGATTT	GAACTCAGGT	CTTCCAAGIC
8581	CCATTGATAG	CATCTCCTCA	CAAGTCCCTT	GCCACCCTCA	CGATGCCTTA	GACACTTGCC
8641	TGCCCTTTAT	ACTAAGGAGA	TGCAGGTACA	AGGGGTTTAC	CCATGTAGCA	GCTGAGGCAG
8701	CTGGGGATAG	ATACCAGCAG	CAGGCCTGAT	GTCACCACTC	TAACTCCAGC	ATCCCCAGTC
8761	TGTGTTCCTG	GAGTGTGAAA	ATCCCTACTT	AACAAGATTG	TGCAACAGTC	CTTGGCTCTG
8821	TGACCCATAG	CTGGAAACAG	GATTCTCATT	GATTTGTGGA	ACATGGTGGC	AGCCAGCCAA
8881	AAAGAGGGTC	TGCATACAGA	AGACACGTGT	GGCAAGGCCA	CAGCAGACTC	TGACTACCTT
8941	AGCTTACAGA	ATTACAAGGT	CATAATGTCC	TCTGCTTTGG	TCACCTCATG	TTAAGGACAG
9001	GCCCTAATGA	AGATGGGGCA	GAAGACTGAA	GGAATGGCCA	ACCAATAACT	GGCCCAACTT
9061	GAGACCCATC	CTACAGGCAA	GCATCAATTC	CTGACACTAC	TAATGATACT	CTGTTATGCT
9121	TGCAGACAGA	AGCCTAGCAT	AACTATCCTC	CGAGAGGTCC	ACCCAGCAAC	TGACIGAAAC
9181	AGAAAAAGAT	ATCCACAGGC	AAACAGTGGA	TGGAGGTCAG	GGACTATTAT	GGGAGAGCTG
9241	TGGGAAGGAT	TAAAAACCCT	GAAGGGGATA	GGAACCCCAC	AGGAAGACCA	ACAGAGTCAA
9301	CTAAGAGACC	TGTGGGAGCT	CTCAGAGACT	GAGCCACCAA	CCAAAGAGCA	TACACAGGCC
9361	GGTCCGAGGC	ACCTGGCACG	TGTGAAGCAG	ACATGCAGCT	CAGTCTCCAT	GTAGGTCCTC
9421	CAATAAGCGG	TAGCCTGACT	GCAGTATCCA	ATCCCCAACA	GGGCTGCATA	GTCTGGCCTC
9481	AGTGGGGGAG	GATGCCCCTA	ATCCTGCAGA	GACTTGATGA	GTGGAGAGCT	ATCCAGGGG
9541	AACCCACCCT	CTCTGAGAAG	GGAATGGGGA	TGGGGGAGGG	ACTCTGTGAA	GAGGGGACAA
9601	GGACAAACAA	GAACCTCAAA	TAGGTCAGGC	CCTAAAGGCT	TGCTAAGTAG	CAGIGGCCCA
9661	GCTCTGTCCT	GTTCCTCAGC	CCAAGGCTCA	GCTCCCACCT	GTTTCTGTGT	TTTTCTGGCT
9721	TTTCATGGGC	CTAGGACTTG	GTGACCAGTT	CAAACAATGG	GGCCTGTGGA	AGACACAATA
9781	TACAAGACTA	GGGACATICC	TGTTCTGCTG	ACTATCCATA	GCCTGATGTA	GGIGGAAGGA
9841	CCCAATCACT	GGATTTCTAC	CCTTGCACAA	CCTTGACAGC	TGAGGGCCTC	TCAGAAACCT
9901	ATTTCTTCCA	CIGAAAAATG	AGACTCTCAA	ATGAACGTCG	TGACAATCAT	CAGGCIIAII
9961	. AAAGAGGTGT	ATCTAACCTG	AATGGCAAGC	AGACAGCAGG	CAAATGTCTG	TATEMACCIC
10021	TAGGAAGGAC	AAGAACTGCT	CACTGCTGCC	CCCCAGGAGG	CCATTIGCIG	AMACAGCIGC
10081	TCTCCTGCTG	GTGCACAGGC	CCTGCCTTCT	CATTGCAGCC	ACAGCCCCTI	TCATCCTAAA
10141	CCTCCTGTCA	GGTCACTGGG	AAACAGATCA	AGATGGAACA	GGACAGCICC	ACACTETGAA
10201	TAAAAAAACAG	TGGTCATGGC	TATTCATAGG	GGTTTATGCT	CCTTCTGAAA	TTCACAGTGG
10261	GAGCTGTGGG	CATGAACCAC	AGIGITCGAG	CTUCHCIICO	CACACAGTCC	TGCTGCTTCT
1032	L GGTGAGCTCA	GTAAATGTGA TCCTGAGGAG	GCIGGAGGI	, WCICGIOWAY	DEDGGCCACA	GTCTAGTTGA
1038	L GTTCCCAATA	C ACCAGTTTCT	ACGACACA1	. INCITIOTIC	TETETETETE	TGTGTGTGTG
1044	L CCTGAGAGIT	GTGTGAGTGG	ACCICCACAC	, vacvaycecta	ACACGTTGAG	GTCAGAGGAT
1050	1 TGTTGTTCG	GIGIGAGIGC	MOGIGCACA:	L COLCGENCAC	TGGAGAACAA	ACATGGGTCC
1056	1 AACTATCAG	GGAGCAAGTC	CONCURC	COLOGNACIO	GCTCACATAC	ATTTTACCTA
1062	1 TTATTCCAG	GGAGCAAGIC	י אכינעדונטכי	ר ייאטרערנינדר	CCGATGGCAA	TGCACCACCT
1068	1 GACAATGGA	C ATAGGAGTTO	COTOTOTO	CACCCCACA	CCCCTTCAC	CAAAACGTTT
1074	CICIACCCA	TATCTGGTA	ACTOCACIO	A CAROCCARA	CAGTATTAAG	AACATGGAAT
1080	1 TCAGTTACT	A CCTGGATCT	CONTRACTO	C ACCCTAGATO	GAGTTGCTGA	GTTTTCACCT
1086	1 CALLIGGGA	A TICCCCCCTI	CCTTCTATG	TTTATTCTGA	AACCAGGGGA	ACTCGATTCC
1092	T CARACTERIA	A TICCCCCCIA		T GAATTCACAT	GTCATCTACT	GCTAATCCAT
1098	1 TCCCTTTGG	C TOCOTOS CS	2 AGACACACT	A CAGTCATGG	CAATGTCAAG	GTAGGACAGA
1104	1 TGGIAGIAI	E TOOCICACA	CTCCTCTTT	T CATGACTAAC	CCTCCTCAGO	ACAGTGACCA
1110	1 IGIGAAICA	T 1000000000000000000000000000000000000	· OTOCIOIII	T AGAATTGCTO	GAATITICIA	TITTGAGAAA
1170	T TOWACCTAC	4 GGGGVGGV	דעעעעטעעע יי	C ATCTAGAAAC	CTGGTTTAA	ATACAGATGG
1122	I TAATAGCCT	T CARROCAL	CCS SACALANA	TATTGGCCC	TCACAGAGG	TGGCTCACTC
1128	T TIGHTTCAG	T CHAMBANGTO	י האולניטעער י היינייייייייייייייייייייייייייייייייי	G GGTCAGGTG	ATAGGAAAG	TNGTCTGGGA
1134	L CAGCAGAGG	C CYCNYLLLANC	CILOGACAC	C TGTTGGCTT	r TITITITITI	AATGAGTTCT
1140	L CAUTGAGAA	CHCWWIIGH	T TACKCASAA	מ כדדרהמהררי	A ACCCAACAG	A ACATTCTTCC
1146	L CAAAAAATG	W GIGGCINGC	L CALAMANATY T TVOOCANYT	C DATACTGAD	C TGACCCCTT	TTGGCAGTCT
1152	ATTGATICA	C NACHENCIC	L Jahray Valla	G GCACAAGGT	CATCACTGT	CACCCAAGCAT
1158	I CAAGITIGA	C CANCHINGGO	C LIIOMACII	G GCTAACCTT	A GATATTTC	T TTGGAGTATC
1164	LI CCIGGGIGA	C CITICONIA	y GGGGALACTI	T CTCATCCTC	G GACCACAAT	A TAAGTCACCC
1170	JI TILAGAACA	I CCAGGAAAI	G CPGPGFFF	A AGGATETET	C TCGTGTGCA	CCTTCTTCAA
11/6	1 INGANICCE	r cononicoi	56/58			
			>D/>X			

11821	AGCAGTGAGT	AGTGACTCCA	CTAAACTGAG	TTCCCATCTG	AGAGTCCACA	GGAGGCTTTG
11881	GGGCAAGAAG	CAGAGGGAAG	GCACTGTTTG	TGTTGGTAAA	GTTTTGACTC	TAACAAATTT
11941	GAAGACATAG	ATGACATTGT	GTCAGACTAA	CAACAACCTA	GACTCATGTG	GGTTCTGTTT
12001	AGGGATCAGA	TTTTATTCAT	CAATGACTTG	TCTTAGTGTA	TAGAGAAAGG	CTTCCTACTG
					CCTAGGGACT	
					CTCTGGGGAG	
					AAACAGAATT	•
					ATGTACATGA	
					AGAAAATAGG	
					TTTCCTGAAA	
					GGGCCCTACT	
					TGTGAAGAGA	
					GTGGGTTTTT	
					AAAGCCGCCA	
					CCGCAATCTG	
					AGTCGTCATG	
					GACGCATGGC	
					ATCGCCTATC	
					AAGGGCTGCC	
					AAAACAGGCG	
					CAGGCAGATA	
					CTGGATAACG	
					TTTACCGGGC	
13201	GGCGTTTCGC	GAACTGGAAC	GGGTGCTTAA	TTTTCCGCAA	TCAAACTTGT	GCCTTAAACG
					CCATCGGAAC	
13321	CGCCGACAAT	GTCTCATTAA	CCGGTGCGGT	AAGCCTCGCA	TCAATGCTGA	CGGAGATATT
13381	TCTCCTGCAA	CAAGCACAGG	GAATGCCGGA	GCCGGGGTGG	GGAAGGATCA	CCGATTCACA
13441	CCAGTGGAAC	ACCTTGCTAA	GTTTGCATAA	CGCGCAATTT	TATTTGCTAC	AACGCACGCC
13501	AGAGGTTGCC	CGCAGCCGCG	CCACCCCGTT	ATTAGATTTG	ATCAAGACAG	CGTTGACGCC
13561	CCATCCACCG	CAAAAACAGG	CGTATGGTGT	GACATTACCC	ACTICAGIGC	TGTTTATCGC
					GAGCTCAACT	
					TTTGAACGCT	
					CAGACTTTAC	
					GTGAAACTGA	
					GGTTTTACGC	
					GGATCACAAC	
					TACTCTCCAG	
					ATCCACACAT	
					GTAGTCTCCA	
					TGACTTGCCT	
					CTTTAATTCC	
					GTCTACAGAG	
					САВАВАВАВА	
14401	GTTGGCAGAG	TGTGGGTTAT	ATACCAGGTG	GAGATTICAA	ATGAGTGGCT	GAAGCIGIAG
					ATGTAGAGAG	
						CAGCCAGCTA
						AACAGATAGA
					TAAGGGTCAT	
					TATCACCTCC	
						CTTTGTCATA
					TGATGCCAAA	
						ATCAGGGTCC
						AACACTGTCA
						ACAATTAAGG
						AGATTCATGA
						GACAGAATIG
						CAGGGACAGT
15241	LAACTGTAGG	CCAGTCCTTC		TGGGTTTTAT	GGTTTGAATC	_TGCAAAGCCT
			57/50			

· · · · · · ·

Figure 23 (continued):

//

```
15301 TCCATGTGCT CAAAGGTTTG AACATGGAGC CTCCTCCTGG TAACACTGTA TTGGAGGCTT
15361 TTGAGACTGG ATGCTCTTTG GTCCCATGTT TTGCTACATC ATCTGTCAAG ATATGACCCA
15421 GGCATGCTAC CAGCTACCAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
15481 TAGACTIGTA TCTCCTAAAA ATGGAATCAA AGCAAACTTT TCCTGCATTA AGTTTTTTTT
15541 TTTCTGTTAA GTGTTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
15601 GAGTTGAGGT TCATTGCTCT AGCAAGTTGG ATCAAATTTT TAGGGCTTTG GAACTGATTT
15661 ATAAGAGACA TGTAGAAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCACC AGTTTTTAAG
15721 AATAGTTTAA TACACCATGG GAATTGTGAA AATCAGAATG CTCACACAAA GGCAGACAGG
15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
15841 GAAGCCATTC GGCTACGTTA GGGAACGTGT GTGGCTGTGC TTGGCCCATG CCCTGGCAAT
15901 CTGAATGAGG CCAAATTTA AAGGAGTGGA CTAACTCGAT TGTCAGAGAA AATATCAAGA
15961 CAGACCACCA CTCAGGCTAT GCCGTGTTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTTAG TGAAGTAAGG
16081 GGTGTGGGTC CCTAAGTGGA TGGTGCATAA ATCTATGTAG GTGATGCCTA AGTGACACTT
16141 GATAATCCAA AATATCAGCA ATGTGGAATG TCTTCCAAGG AGACCTGTAG ACACACATTT
16201 TAGAACTTIG CTCATGGCTG TAATAAATAG CTAGCTAGAA ATCATITCCT GAAGAGGTTA
16261 GTCTGAGTTA CGGTTCCAGG GCAAACATTC AGTGATGGCA AGGAAGGCAT TGCAGTCAGG
16321 AGCCAAAGGT CAGCTGGTCA CATTGCATCA AGAGTAGAGA GTCAGAGTGT GAGTAGAAAG
16381 AGGATACAGG TTATAAAACC TCACTGTCCA CTCTCAGCAA TCCATTTTCT CCTAAAAGGC
16441 TITACCTTCT AAAGATTITA GTCTTCAAAA CCAGTACCAG TAGCCTGGGA ACAAAAGTTG
16501 AAACAAATGA GCCTTTGTGG GGCATTTCAC ACTTAAAACA GGGCATCACC TAGGAGGAGC
16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAGGAATG CTCAGGCTAA TAAGGGGTCC
16621 TCTATCTGAG GGACCCTATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGTAAATGGA
16681 TGCTACCAAT TTGACATTTG TAGACCTGCT ATTGTGTGCT TCTTTATTGG GCTCTCCCAT
 16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
 16801 AGGAGAGAG GAAGGTTAGA AGAGAAAGTG GGTATAGATC TATTTAGACT ACTTCCTGCT
 16861 GATTAGGGGC AAGTCCAATC GTCATTGTCA GGATACCTCC AACCAGCAAC CAGCAAACCA
 16921 GCAAATCAGA AACAGCAAAA GCAGCCAACA AGGCAGCACT AACCAGCAGG ATTGGGGTCG
 16981 GTAGCGTGGG AGCAGTCACT ACTGGTCTTC TCATGGCTTT GGCATTAATA CTCTCTCAAG
 17041 AAATTCCGTA ATTTTTTCCC CACCACCTGA AATTCCGTAA TTTTAAATGC AAACTATCTA
 17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATTT
 17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAA ACAAAAACAT ATTCACATCA
 17221 CATAACTGTT TITTTTTCC AATTITTTAT TAGGTATTTT CTTTATTTAC ATTTCAAATG
 17281 CTATCCCGAA AGTCCCCTAT ACCCTCCCAC CTCCCTGCTC CCCTACACAC CCACTCCCAC
 17341 TTTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCCT
 17401 CTCTTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
 17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTCA
 17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGGCTC TGTGTTTTAT CTAATAGATG
 17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
 17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT
 17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GC
```

SEQUENCE LISTING

```
<110> University of Guelph
     Forsberg, Cecil W.
     Golovan, Sergeui
     Phillips, John P.
<120> Transgenic Animals Expressing Salivary Proteins
<130> U Guelph - 03
<140>
<141>
<150> 60/130,508
<151>. 1999-04-23
<160> 7
<170> PatentIn Ver. 2.1
<210> 1
<211> 20623
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Lama2/APPA
      plasmid
 <400> 1
 tcgagagtat ctttgtcagc tgtgcctcca acaaaggggt actgttgccc acatagaaag 60
 atctaaacta attaattaat ccctcacccg caaatctttc agtcactaag ttagcacgat 120
 tgttgaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180
 ctgaggagac acctgcatct gactaagaag agccacggtg ttagttgaat ggtgtggagt 240
 agggtggttc tgtgggacag tagaaaatcg agaggcatgt gccgtttagt gaactgatgg 300
 aagctacccc aaacgacaga gattgtcagt caggccaatc cgtttcgagt ttgatgggca 360
 gccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420
 agggattgag agaccetgae aggegeaagg ecetaacaca cacacetace aceteacttg 480
 acaaagctgc caaagaccaa agacttgttc tccattagaa atgacagctg gcttgacccg 540
 acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgtttc attcagtatt 600
 ataaaaggac agtattacag attttgttgt acactgctgt tacatgtggg gcagtgtgtc 660
 tttaagtagg gtaaagtact ctttaaaaat gggtcctaga tatttttcc tttaactcaa 720
 gtctcttact gtttaaatga tttttatttt gtttaatatg gaggaaaaag aagcgtaaat 780
 ggacaatata tatttagaga aagatggtta gctgtcagaa aaatatgcaa atcaaaatca 840
  caccaagact gcagcacacc cctgtcagat ggctgtgatc aagaaaataa atgacaatga 900
  cacactggag caaccactgt ggaaatcagt atgaatggtc ctcaaaaacc tgaagataga 1020
```

gcggggcgtg	gtggcataca	cttttattcc	cagcactggg	gaggcagagg	caggtggatc 1	1080
tctgagttcc	aggccagcct	ggtctatagc	acaggttcta	ggacagccag	ggctacacag	L140
aaaaaccctq	ccttgattaa	accaaaccaa	accaaaccaa	accaaaccaa	accaaaccaa	1200
accaaaccaa	accaaaccag	accaaaccaa	aacactgaag	atagaacttc	agtattccat :	1260
tectagatat	atacccaatg	gagactaagt	cagcaagaca	cctgcacagc	catgttcact :	1320
actacactqt	tcaccacage	caggetgtgg	aaccagcctg	agtgtccatg	ataaatgaat :	1380
ggataggtaa	ctttcaaggt	aaatggactc	tgctgtgtac	atgcctcaca	ttctgtttat	1440
tcatttttct	ttatgaggtg	tccattcagg	agtcacatgg	tagttctatt	ttcagtcttc	1500
tgaagatact	acactqqtcc	ccacagttta	cacttttatc	agcagtgaat	aagggttcct	1560
ctateettae	catcatttqt	tqtaattttt	cttgatgacc	ctctttctga	cagggatagg	1620
atgtaatatc	agtgtgagga	agtacaactt	gttttctaag	tatttattgg	ccccttgcat	1680
ttcttcttt	gaaaactgtc	ggttcctgac	atctgctcag	gtattcattg	gatgttgttt	1740
ctttaatatt	tgagttctta	tgaattctag	atgttaaatc	cctgcctgtg	gttctctccc	1800
attetataga	ctacctcctc	accctggcaa	ttgttgtcct	tgttttgcag	aaacttttga	1860
cttcatggaa	tctcatttqt	cagttttccc	tectetgeta	tagcctgagc	taatgcactg	1920
gtttttacag	agecetggte	tatgccttta	tectectetg	gcagcttcgg	agtttcattt	1980
cttacattta	gatetttgat	ccactttgaa	caagttttgg	agcagggtga	gagatacgaa	2040
tctagttcca	ttcttccata	tgtgatccta	gtttacatag	catcgttggt	tgaagaggtt	2100
ttattttatt	tttaaataat	gtgtcataaa	aaacgaggtg	gttgtagcag	tgtggatttg	2160
tttctttatc	ctttgatcta	caggtcttgt	tttgtgtcag	tctcatgatg	ttttattgct	2220
atggctctgt	catacagtet	gaggtcaggt	attgtgatat	accttcagta	ttgctccctc	2280
agactcaggt	ttactttagc	caggagtcat	cttactcagt	gctcttagag	ctcccccagc	2340
atgtagetge	tactattctt	agttgataaa	tcaggaaact	ggggctcaga	gagattaact	2400
gtcttgaact	acttctqqqq	aggtgaaacg	tggagacact	aaactgtgtt	taccctgtac	2460
toctccagta	actatcaagt	gctgggctac	agcaaagca	ctatactata	tattactcag	2520
gaggtggaaa	aactcagcct	cccttggggt	tcccaaget	c ccaggtgtcc	agtcactgct	2580
ggaaacctca	a tqqagtctga	aaggaagggt	tgagggtaca	a tggggcagcg	atgaggagcc	2640
tagaactaga	atctcccaaa	cacctggata	tecagatge	c actgggtcag	ggggagttgg	2700
gaacagagtt	t gggatgtcca	tggacctgtg	acaaggccag	g ggccaggggg	g aggataactc	2760
togetttact	t aatttgcgaa	agtccttago	ttagcagcag	g ttgtctggga	a gcacagaggg	2820
geettetgt	a agaggeteag	gcagtgccg	tctgtaggc	g aaggtette	ccatgttccc	2880
catggtggt	t cttgatgaa	a gagacagtco	ttggctcca	a actggttta	tgattgttca	2940
ttqtqqaaa	a tgggtgcac	a ccaccttctc	agggtggac	c agagatcaa	a taccttttgc	3000
agggaggaa	t atctgggaa	g ggacgcttad	tggctaaac	c ctcagggcc	t ctagatacat	3060
cattagcat	g gagaactct	g ttctgggcta	a catgaccac	a ggccacatt	t ccacaagcca	3120
catqtqqqa	a gtgtggcac	a tgttctagge	c caggaatct	g gtagggagc	g tggagccacc	3180
taccatccc	a ggtgggtgc	c tgggtgcca	g ggaccctga	a ccegctcaa	c cttaccaagt	3240
tteetggea	q qqtccactg	t cctacacag	a agctggagg	a ggtgtgagg	g ttgtgtcttt	3300
gtggaatgt	c ccatgctgc	t tggggctca	g tttctccac	c tgtacctca	t tggtttgggt	3360
ataaaaaqt	g gggatactt	t attattctc	t gacteggte	c tgaggaaaa	a gcatcgtggc	3420
agtccagga	a ccacaccct	g aggtteetg	c actgaaggg	ja ctccctaag	t ctctggagto	3480
teteceett	c acagagets	c caaagtcta	g gttcttttg	ya ggataacag	a gccatgcttg	3540
gtaagcaga	c aacagcatt	t gtttactca	a ccttcttt	g tcagctccc	t cttcataaac	3600
aagttgaga	ac accatocto	g cttgaggaa	g acttctaaa	ng ccagacaac	t gtgcaaggaa	3660
gaagaagaa	ag gggcaagt	g agttageet	g gatgtagco	c tcaaagtct	c cagagacca	3720
ccatgaagg	ac tcaaqtqq	g ggcaagacc	t gcagcagco	ea agcatctgg	c aggagagga	3780
cctaggaaa	c cctctacca	t gacacacat	t cttcctgca	ag gtcacactt	a ataggccat	t 3840
tettattt	o atctatcat	g gtgttctgt	g cgagattaa	at gaggtgtta	t gctgcgaac	a 3900

			**	h	atoggogooo	3960
					gtagcagaaa	
					aagccctcca	
					gctacagctc	
					cttgtttggg	
					aggccctcag	
					acaccctggt	
					ccacggcttt	
					tagctatgat	
					gctcacggag	
atgctaggga	agaaacgtct	tctaactact	gaggttacta	agttcctggt	ggttgtctct	4500
gcctttccct	tgttaaagtc	accttgaagt	tagtgcagaa	gaaatcagag	cccagtcaca	4560
gagtaaatat	ggtcctgaag	atttcctttg	agtgcccaga	atccatgaca	tttcaagagc	4620
cctctttgta	ccttaagtca	tttggggttg	tatcttctgc	ttgatgtatg	tgtgtgtgtt	4680
					atttgcaatt	
					cagaaaagag	
					attccctcct	
					cttgggatgg	
					tgggtctcat	
					ttttgctatg	
					gacagctgtg	
					gaagagaaag	
					tatcacatat	
					ggggatatct	
					tctggctcaa	
					gcactttatt	
					aattatcaca	
					ccctcttcca	
					aagaaaagat	
					tcataacaat	
					ataagcacga	
					aggttacaaa	
					ttgaagagaa	
					acttgagccc	
					cctagacacg	
					acttccacct	
					cctctgtgta	
					tttgtctttc	
					atcttcctcc	
					gagctagctg	
					a gaacttagat	
					tatcaacagg	
					a cagatggatg	
					g caaaggacag	
					c aactgaccac	
					c agccaagggt	
					g tccagccttt	
					g gatactgggt	
ccagccttt	g gtaccttca	c tctcagggta	a ccccatcact	gtgtcttgg	a gagcctaggo	6780

accetogote	cageetteag	tacctgcgct	ctcaggacac	cccaccattg	tetettgeec 6	840
catctcttct	tectettect	ccctttcatt	gtctcttctc	tgtttctttc	ttgactctcc 6	900
trtcccctca	caccctcact	ctaqttctcc	ccttccctct	ctgcatcacc	ctattctctc 6	960
tataataat	ccactttcct	ttatctctca	tgcttctctc	ctccctcaaa	tacttgtcac 7	7020
ccactatact	tcaggggcca	gctctagtga	caaagctgtt	aatagcaaga	ctctcagatc 3	7080
tocaacggct	cagaggagcc	agacccacca	agaactctct	ccaggtccaa	tttcaggttc 1	7140
cttcgaaagc	tttcagcaaa	tactcaggga	acatqccact	aacaagaaga	tgcaaattcc	7200
agttgagagt	gggaaaggg	cttacataga	tcccatcttc	caggccaagg	tcagagggc '	7260
tetatataat	ccggattgac	aggeteaga	acaatqtttt	gtttttaagg	tttatttatt '	7320
traggigetta	atatetttae	ttgcatgacc	ttatqtqcat	catgtgtgtg	caggttcctg '	7380
atgacagtag	aggagggett	tgaatccctg	qqqataggaa	gttacaggaa	attataagct '	7440
actttataaa	tettetaget	ttcccaacaq	aagtgaatgc	tcttcaccac	tgagccatct	7500
ctctagggg	aagagacatt	gctttatgga	tataattgtg	tgtgtgtgtc	aacattgagg	7560
aaaggaaat	алалалала	cttcagccgc	taaggttgta	cagtttcact	aattgctact	7620
tttagttata	ataaaatggc	aggtgcttca	acatttatat	atacaaaaac	ttccctgctg	7680
gtggttcaac	totgagaact	ggggtaagtg	ggtgagttct	cttttctgt	ctctgtctct	7740
gtgtctctcc	ttccattctt	tettaaaqqa	aataaacatt	gcagctgggt	tatagctcat	7800
caatatogaa	gttacagaag	tgaaaaaagg	cattgccttg	gtgggtggtg	ttaccagctg	7860
atttttggtt	gtcctgcaag	qaqqtctggg	gactggctgc	tetgtetetg	tctgtatgag	7920
tgagggaagt	ctqqqqaqca	gattccctaa	ccttcagcct	ggcctggttc	ctgagtgaac	7980
ccacctctc	togtectagt	agctttttcc	aaacaggaat	ctgagtggtg	acagggaaca	8040
agtaccagcc	cattgcttaa	gtgccagggt	tagtgagggc	aggaagctgc	catagctggg	8100
attagtagtt	gtattggatg	taggaagtcc	tatcctggga	cagctaatcc	ttaatgcttc	8160
actggagatt	ttcaatgaga	aatttatccc	acggcccata	tggccccatc	cttttgtctc	8220
caacagccaa	gtattttcca	ttagaggaga	cttcctgtac	acttgatgga	tgctcattcc	8280
aaggtgactt	ggggcagtca	gtacagactt	gggatgacct	ctgacagcct	aacctctccc	8340
caacaagggc	cctctatgtt	tgctatgtaa	tgtaatgtca	gacattgtca	ggagtgtccg	8400
cagcacagcc	tgcccagtgt	gagggctctc	ataggtttcc	cactgtctta	tctacacagg	8460
gataacgagg	aggtaagctg	cagttcccag	tctcacttca	cagaggaaga	gataacccca	8520
tcccaggtca	tgtagccago	agtggaaaga	atgaggattt	gaactcaggt	cttccaagtc	8580
ccattgatag	catctcctca	caagtccctt	gccaccctca	cgatgcctta	gacacttgcc	8640
tgccctttat	actaaggaga	tgcaggtaca	aggggtttad	ccatgtagca	gctgaggcag	8700
ctggggatag	ataccagcag	caggcctgat	gtcaccacto	taactccago	atccccagtc	8760
tgtgttcctg	gagtgtgaaa	atccctactt	aacaagatt	g tgcaacagto	cttggctctg	8820
tgacccatag	g ctggaaacag	gattctcatt	: gatttgtgga	a acatggtgg	agccagccaa	8880
aaagagggto	: tgcatacaga	agacacgtgt	ggcaaggcca	a cagcagacto	tgactacctt	8940
agcttacaga	a attacaaggt	cataatgtc	totgotttg	g tcacctcate	g ttaaggacag	9000
gccctaatga	a agatggggca	gaagactgaa	a ggaatggcc	a accaataac	ggcccaactt	9060
gagacccato	e ctacaggcaa	gcatcaatte	cctgacacta	c taatgatac	t ctgttatgct	9120
tgcagacaga	a agcctagcat	aactatcct	c cgagaggtc	c acccagcaa	c tgactgaaac	9180
agaaaaagat	t atccacagge	aaacagtgg	a tggaggtca	g ggactatta	t gggagagctg	9240
tgggaaggat	t taaaaaccc	t gaaggggat	a ggaacccca	c aggaagacc	a acagagtcaa	9300
ctaagagac	c tgtgggagc	t ctcagagac	t gagccacca	a ccaaagagc	a tacacaggco	9360
ggtccgagg	c acctggcac	g tgtgaagca	g acatgcagc	t cagtctcca	t gtaggtcctc	9420
caataagcg	g tagcctgac	t gcagtatcc	a atccccaac	a gggctgcat	a gtctggcctc	9480
agtggggga	g gatgcccct	a atcctgcag	a gacttgatg	a gtggagagc	t atccaggggg	9540
aacccaccc	t ctctgagaa	g ggaatgggg	a tgggggagg	g actetgtga	a gaggggacaa	9600
ggacaaaca	a gaacctcaa	a taggtcagg	c cctaaaggc	t tgctaagta	g cagtggccca	9660

4

getetgteet gtteeteage ceaaggetea geteecacet gtttetgtgt ttttetgget 9720 tttcatgggc ctaggacttg gtgaccagtt caaacaatgg ggcctgtgga agacacaata 9780 tacaagacta gggacattcc tgttctgctg actatccata gcctgatgta ggtggaagga 9840 cccaatcact ggatttctac ccttgcacaa ccttgacagc tgagggcctc tcagaaacct 9900 atttetteca etgaaaaatg agaeteteaa atgaacgteg tgacaateat eaggettatt 9960 aaagaggtgt atctaacctg aatggcaagc agacagcagg caaatgtctg tatcaacctc 10020 taggaaggac aagaactgct cactgctgcc ccccaggagg ccatttgctg aaacagctgc 10080 tetectgetg gtgeacagge cetgeettet cattgeagee acageceett cetgtetgaa 10140 cctcctgtca ggtcactggg aaacagatca agatggaaca ggacagctcc tgatggtaaa 10200 taaaaaacag tggtcatggc tattcatagg ggtttatgct tcttcagtcc acactgtgaa 10260 gagetgtggg catgaaccac agtgttcgag gtagagttgg ggttctgaaa ttcacagtgg 10320 ggtgagetea gtaaatgtga getggaggte actegtgaga cacacagtee tgetgettet 10380 gttcccaata tcctgaggag acgacatc tactttgttc agaggccaca gtctagttga 10440 tgttgttcgt gtgtgagtgc aggtgcacat atgatagcgt acacgttgag gtcagaggat 10560 aactatcagg cgttgtcccc tcctactttt cctcggactc tggagaacaa acatgggtcc 10620 ttattccagg ggagcaagtc gctgttggct gacacatctt gctcacatac attttaccta 10680 gacaatggag cetecateag agtattaett tageteetea eegatggeaa tgeaceaect 10740 ctctacccac ataggagttg ggtctccaca caccccaca cccccttcac caaaacgttt 10800 tcagttactt tatctggtaa agttcatcag agaatgaagc cagtattaag aacatggaat 10860 catttgggaa cctggatcta gcaatacccc accctagatg gagttgctga gttttcacct 10920 cagattataa ttccccccta gcttctatgg tttattctga aaccagggga actcgattcc 10980 tccctttgga ccacagacat cctggcttgt. gaattcacat gtcatctact gctaatccat 11040 tggtagtatg tggctcacag agacacacta cagtcatggc caatgtcaag gtaggacaga 11100 tgtgaatcat tcccccagtc ctgctgtttt catgactaac cctcctcagc acagtgacca 11160 tgaacctact tttcccctcc ttttattttt agaattgctg gaattttcta ttttgagaaa 11220 taatageett gggeageatt aaacaaaate atetagaaag etggtttaaa atacagatgg 11280 ttgagtcagt gaaagagtga ggaatgtcat tattggcccc tcacagaggc tggctcactc 11340 cagcagaggt ggttgaagct cttggacacg ggtcaggtgc ataggaaagg tngtctggga 11400 cactgagaac cacaattgaa caaacagaac tgttggcttt tttttttta aatgagttct 11460 caaaaaatga ctggctagct taggcaaata cttcgagcca acccaacaga acattcttcc 11520 attgattcat tetggatett etttetagae aatactgaae tgacceettg ttggcagtet 11580 caagtttgac aacatagggc tttgaacttg gcacaaggtc catcactgtc acccaagcat 11640 cctgggtgac ctttgggttg gaatatcttg gctaacctta gatattttct ttggagtatc 11700 tttagaacat ccaggaaata gggcttgatt ctcatcctgg gaccacaata taagtcaccc 11760 tagaateeca ggagategtg cagagaaaca aggatetete tegtgtgeat cettetteaa 11820 agcagtgagt agtgactcca ctaaactgag ttcccatctg agagtccaca ggaggctttg 11880 gggcaagaag cagagggaag gcactgtttg tgttggtaaa gttttgactc taacaaattt 11940 gaagacatag atgacattgt gtcagactaa caacaaccta gactcatgtg ggttctgttt 12000 agggatcaga ttttattcat caatgacttg tettagtgta tagagaaagg etteetaetg 12060 gagtgtaggc tcaataatga cagaagagat agctatttcc cctagggact gtgctgctcc 12120 aagtttggtg gagaaaggca gtggggaacc tagatgtgct ctctggggag ggggtctgaa 12180 getggettea tagaaggtgt gaagttttge tgaaacatet aaacagaatt atagettagg 12240 aaagtgagca ggcaaggcag ggaatgtgtt gcatatgtat atgtacatga atatattatg 12300 ttatagatac acacacattt gaacctcatt tgcagatgac agaaaatagg ttattttgcc 12360 tetettaaet getaageaca atgaetteea giteeateea titeetgaaa tgeeacaatt 12420 teatttttea ttgtggetga ataaaattee attgeagaet gggeeetaet teateeaete 12480 ctgagggcag gcatatcccc tggctccatt tcttacctat tgtgaagaga agtgcaactg 12540

tcttgttgaa aggcaagcgt gagagaggca ggcactaatt gtgggttttt gtttcttctt 12600 cctgctatga ctctccattt gtcagaacca aagatcgata aaagccgcca ccatgaaagc 12660 catettaate ccatttttat etettetgat teegttaace eegcaatetg cattegetea 12720 gagtgagccg gagctgaagc tggaaagtgt ggtgattgtc agtcgtcatg gtgtgcgtgc 12780 tccaaccaag gccacgcaac tgatgcagga tgtcacccca gacgcatggc caacctggcc 12840 ggtaaaactg ggttggctga caccgcgcgg tggtgagcta atcgcctatc tcggacatta 12900 ccaacgccag cgtctggtag ccgacggatt gctggcgaaa aagggctgcc cgcagtctgg 12960 traggtregeg attattgctg atgtregarga gregtarcegt aaaacaggreg aagrettrege 13020 cgccgggctg gcacctgact gtgcaataac cgtacatacc caggcagata cgtccagtcc 13080 cgatccgtta tttaatcctc taaaaactgg cgtttgccaa ctggataacg cgaacgtgac 13140 tgacgcgatc ctcagcaggg caggagggtc aattgctgac tttaccgggc atcggcaaac 13200 ggcgtttcgc gaactggaac gggtgcttaa ttttccgcaa tcaaacttgt gccttaaacg 13260 tgagaaacag gacgaaagct gttcattaac gcaggcatta ccatcggaac tcaaggtgag 13320 cgccgacaat gtctcattaa ccggtgcggt aagcctcgca tcaatgctga cggagatatt 13380 teteetgeaa caageacagg gaatgeegga geeggggtgg ggaaggatea eegatteaca 13440 ccagtggaac accttgctaa gtttgcataa cgcgcaattt tatttgctac aacgcacgcc 13500 agaggttgcc cgcagccgcg ccaccccgtt attagatttg atcaagacag cgttgacgcc 13560 ccatccaccg caaaaacagg cgtatggtgt gacattaccc acttcagtgc tgtttatcgc 13620 eggacacgat actaatetgg caaatetegg eggegeaetg gageteaact ggacgettee 13680 cggtcagccg gataacacgc cgccaggtgg tgaactggtg tttgaacgct ggcgtcggct 13740 aagcgataac agccagtgga ttcaggtttc gctggtcttc cagactttac agcagatgcg 13800 tgataaaacg ccgctgtcat taaatacgcc gcccggagag gtgaaactga ccctggcagg 13860 atgtgaagag cgaaatgcgc agggcatgtg ttcgttggca ggttttacgc aaatcgtgaa 13920 tgaagcacgc ataccegett gcagtttgta aggtaccegg ggatcacaac ttgccctctg 13980 aagaggaaga acagaaggat gccacaactc tcctgctggc tactctccag tggtttcatc 14040 ttacttctga tggcatttcc ctctagaaag tgctactatc atccacacat ttctacctga 14100 gaccacccaa aggaccctcc caaattctct teetetetga gtagteteca caeetgttac 14160 caccatecca gaattaaaat eetaaetgea etetggegtg tgaettgeet eagteettge 14220 aataagagtt gttggcagtg ccaggcgtgg tggcgcacgc ctttaattcc agcacttggg 14280 aggcagaggc aggcggattt ctgagttcga ggccagcctg gtctacagag tgagttccag 14340 gacagccagg gctatacaga gaaaccctgt gtcgaaaaac caaaaaaaaa aaaaaaagtt 14400 gttggcagag tgtgggttat ataccaggtg gagatttcaa atgagtggct gaagctgtag 14460 ccagaaggaa cttagaggat agctcataac ttaaaaagaa atgtagagag tagcagaaac 14520 attgagagag tgggcacaca gccactgtgt gaatgtggca gaacacaatc cagccagcta 14580 tacatgcata agtgtatatt ggcgccatcc tgactgatga gacacaggaa aacagataga 14640 cggggttagg tggccatggc ctttcctgcc tgcctcttcc taagggtcat ctcaagacct 14700 tatgetetet taaetettee attgetaett agettetaga tateaeetee agattagtet 14760 ccttgggtac atcagtgatc ctggtgatat ccagggcttc ctgattccat ctttgtcata 14820 gaggetgeaa etaaagaggt ettettaata etteaeaeee tgatgeeaaa aggaagaeae 14880 agaagttcac agaggtgaag tgattcatgt aggacataca gtgagcaagc atcagggtcc 14940 ggattatctg actctactct aacttttatg taaatgtgct ttatgccatt aacactgtca 15000 ttcctgtgct tcagctctgg gagactccca agcactctta ggcacaagcc acaattaagg 15060 gactetgaca etetgeattg attaattage atggtggtet etatgtttee agatteatga 15120 ttgtttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatcc acctetetea ggaageetet atetggaaaa gettacaact cagggacagt 15240 aactgtaggc ccagtccttg gtgtccaaaa tgggttttat ggtttgaatc tgcaaagcct 15300 tecatgtget caaaggtttg aacatggage etceteetgg taacactgta ttggaggett 15360 ttgagactgg atgetetttg gteccatgtt ttgetaeate atetgteaag atatgaceca 15420

			tatagagatt	tastattata	cocacataa	15480
		agactatgcc				
-		atggaatcaa				
_		acagggacaa				
		agcaagttgg				
-		tctgaagctg				•
_		gaattgtgaa				
		gtgagagggc				
		gggaacgtgt				
ctgaatgagg	ccaaatttta	aaggagtgga	ctaactcgat	tgtcagagaa	aatatcaaga	15960
cagaccacca	ctcaggctat	gccgtgtttg	tgaccgacca	gctactctta	gccagctcta	16020
ttgtgaaatt	ccagagcaat	tatcagagca	tgaagataca	tacagtttag	tgaagtaagg	16080
ggtgtgggtc	cctaagtgga	tggtgcataa	atctatgtag	gtgatgccta	agtgacactt	16140
gataatccaa	aatatcagca	atgtggaatg	tcttccaagg	agacctgtag	acacacattt	16200
tagaactttg	ctcatggctg	taataaatag	ctagctagaa	atcatttcct	gaagaggtta	16260
		gcaaacattc				
		cattgcatca				
		tcactgtcca				
		gtcttcaaaa				
		ggcatttcac				
		tggcctctgt				
-		aagattcaac				
		tagacctgct				
-		attccacatt				
		agagaaagtg				
		gtcattgtca				
		gcagccaaca				
		actggtcttc				
		caccacctga				
		tctcctagag				
		atcccacacc				
		aatttttat				
		acceteceae				
		cggtactggg				
		ctaagccatc				
		tcatattgtt				
		agctcctcca				
		tatttgacag				
		aaccttgctg				
		tagttctaga				
		taattgcgcg				
		tcacaattcc				
		: ceacaaccec : gagtgagcta				
		: tgtcgtgcca				
		ggcgctcttc				
		cggtatcago				
		gaaagaacat				
		tggcgttttt				
caaaaatcg	a cgctcaagto	agaggtggcg	aaacccgac	a ggactataa	a gataccagg	c 18300

```
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 18360
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 18420
teteagtteg gtgtaggteg ttegeteeaa getgggetgt gtgcaegaac eeccegttea 18480
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 18540
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 18600
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg 18660
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 18720
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 18780
aaaaaaagga totcaagaag atootttgat ottttotacg gggtotgacg otcagtggaa 18840
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 18900
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 18960
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 19020
atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc 19080
tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc 19140
aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc 19200
catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt 19260
gegeaacgtt gttgecattg ctacaggeat egtggtgtca egetegtegt ttggtatgge 19320
ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa 19380
aaaageggtt ageteetteg gteeteegat egttgteaga agtaagttgg eegeagtgtt 19440
atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg 19500
cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc 19560
gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa 19620
agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt 19680
gagatecagt tegatgtaac ecactegtge acceaactga tetteageat ettttaettt 19740
caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag 19800
ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta 19860
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 19920
aggggttccg cgcacatttc cccgaaaagt gccacctaaa ttgtaagcgt taatattttg 19980
ttaaaattcg cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc 20040
ggcaaaatcc cttataaatc aaaagaatag accgagatag ggttgagtgt tgttccagtt 20100
tggaacaaga gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc 20160
tatcagggeg atggcccact acgtgaacca tcaccctaat caagtttttt ggggtegagg 20220
tgccgtaaag cactaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga 20280
aagccggcga acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg 20340
ctggcaagtg tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg 20400
ctacagggeg egteceatte gccatteagg etgegeaact gttgggaagg gegateggtg 20460
egggeetett egetattaeg eeagetggeg aaagggggat gtgetgeaag gegattaagt 20520
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgagcgcgcg 20580
taatacgact cactataggg cgaattgggt accgggccc ccc
                                                                  20623
```

```
<210> 2
<211> 6708
<212> DNA
<213> Artificial Sequence
```

<220>

<223> Description of Artificial Sequence: R15/APPA +

intron plasmid with pBLCAT3 vector

```
<400> 2
ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtcctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtotaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240
catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480
tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660
taaaacatat gtttgaccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720
gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtotoaca aaatgaaaag gaaatatatt cagaaagaga atottgagag aatgtgttgt 840
cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900
taaqataaaq qtaactqtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020
tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140
atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttottaacaa tagaacatat aaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500
gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680
agtatgaatt caaaggtgcc attottctgc ttotagttat aaaggcagtg cttgcttctt 1740
ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgttaacccc 1860
gcaatetgca ttegeteaga gtgageegga getgaagetg gaaagtgtgg tgattgteag 1920
tegtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980
cgcatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
cgcctatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccg cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa 2160
 aacaggogaa goottogoog cogggotggo acctgactgt gcaataaccg tacataccca 2220
 ggcagatacg tecagteceg atcegttatt taateeteta aaaaetggeg tttgccaaet 2280
 ggataacgcg aacgtgactg acgcgatcct cagcagggca ggagggtcaa ttgctgactt 2340
 taccgggcat cggcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc 2400
 aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc 2460
 ateggaacte aaggtgageg cegacaatgt etcattaace ggtgeggtaa geetegeate 2520
 aatgctgacg gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg 2580
 aaggatcacc gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttta 2640
 tttgctacaa cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat 2700
```

	ttaraacaca	atecaceeca	22220000	tatoototoa	cattacccac 2760
					gcgcactgga 2820
					aactggtgtt 2880
					tggtcttcca 2940
					ccggagaggt 3000
					cgttggcagg 3060
					gtataaggca 3120
					ageggatgaa 3180
					acataattgg 3240
					agtgtataat 3300
					ggaactgatg 3360
					gaagaaatgc 3420
	•				aaaaagaaga 3480
					agtcatgctg 3540
					aaagctgcac 3600
					cataacagtt 3660
					gctattaata 3720
					aataaggaat 3780
					tgtagaggtt 3840
					aatgaatgca 3900
_					caatagcatc 3960
					gtccaaactc 4020
					ttcgtaatca 4080
					caacatacga 4140
					cacattaatt 4200
					gcattaatga 4260
					ttcctcgctc 4320
					ctcaaaggcg 4380
					agcaaaaggc 4440
					taggeteege 4500
					cccgacagga 4560
					tgttccgacc 4620
					gctttctcaa 4680
					gggctgtgtg 4740 tcttgagtcc 4800
					gattagcaga 4860
					cggctacact 4920
					aaaaagagtt 4980
					tgtttgcaag 5040
					ttctacgggg 5100
					attatcaaaa 5160
					ctaaagtata 5220
					tatctcagcg 5280
					aactacgata 5340
					acgeteaceg 5400
					aagtggtcct 5460
					agtaagtagt 5520
tegecagtt	a atagtttgcg	g caacgttgt	c gccattgcta	a caggcatcgt	ggtgtcacgc 5580

```
tegtegtttg gtatggette atteagetee ggtteecaac gateaaggeg agttacatga 5640
tececcatgt tgtgcaaaaa ageggttage teetteggte eteegategt tgtcagaagt 5700
aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 5760
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 5820
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 5880
catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 5940
aggatettae egetgttgag atecagtteg atgtaaccea etegtgeacc caactgatet 6000
tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 6060
gcaaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 6120
tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6180
tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 6240
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 6300
cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg 6360
gtcacagett gtctgtaage ggatgeeggg ageagacaag eeegteaggg egegteageg 6420
ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat tgtactgaga 6480
gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg 6540
cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg 6600
ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttg ggtaacgcca 6660
gggttttccc agtcacgacg ttgtaaaacg acggccagtg ccaagctt
                                                                  670B
```

<210> 3
<211> 4060
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: R15/APPA +

<400> 3

intron transgene

ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60 gagagteetg tttggtttaa geaacetetg ttteteataa aeteeataaa aacaggaata 120 ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180 ataggtetaa taacceegaa aatattaeca tgataetgag catttgtaag tateteatag 240 catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300 tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360 tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420 aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480 tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540 ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600 tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660 taaaacatat gtttgaccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720 gattetettt gggtggetge aaattgteea egaataagae aaaatataaa aataaggaet 780 gagteteaca aaatgaaaag gaaatatatt cagaaagaga atettgagag aatgtgttgt 840 cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900 taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960 ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020

tccatggacc	tttgaaatat	aaaatagtca	agcaacttat	caaggaatta	cagattcctt	1080
gatactaaca	caggtaaatc	ccacacgtgt	tttgagacta	catttgctgg	gattttattg	1140
atgtaatagg	tcacatgttt	ttcgggccaa	tgttgctgtt	atteggttac	ttcaagagaa	1200
tagtggcaac	tgatgctatg	tattctaggg	gtttgaagtg	atgtttcatg	attgaaattt	1260
gtaaaagaat	aacatcatca	ttcttaacaa	tagaacatat	aaagtcacac	agaagtgaca	1320
_			tttattacct			
-			actaatccta			
			ttatataacg			
			cacatgattt			
_			gcaacttttc			
			gatttcctgc			
			ttctagttat			
			gcgagatttt			
agccgccacc	atgaaagcca	tcttaatccc	atttttatct	cttctgattc	cgttaacccc	1860
gcaatctgca	ttegetcaga	gtgagccgga	gctgaagctg	gaaagtgtgg	tgattgtcag	1920
tcgtcatggt	gtgcgtgctc	caaccaaggc	cacgcaactg	atgcaggatg	tcaccccaga	1980
cgcatggcca	acctggccgg	taaaactggg	ttggctgaca	ccgcgcggtg	gtgagctaat	2040
cgcctatctc	ggacattacc	aacgccagcg	tctggtagcc	gacggattgc	tggcgaaaaa	2100
gggctgcccg	cagtctggtc	aggtcgcgat	tattgctgat	gtcgacgagc	gtacccgtaa	2160
aacaggcgaa	gccttcgccg	ccgggctggc	acctgactgt	gcaataaccg	tacataccca	2220
			taatcctcta			
			cagcagggca			
			actggaacgg			
			cgaaagctgt			
			ctcattaacc			
			agcacaggga			
			cttgctaagt			
			cageegee			
			aaaacaggcg			
			taatctggca			
			taacacgccg			
			ccagtggatt			
			gctgtcatta			
			aaatgcgcag			
			accegettge			
			cgcctgaata			
			aaccttactt			
			aggtaaatat			
			tattttagat			
			aggaaaacct			
			ctcaacatto			
			cagaattgct			
			ctatttacac			
			ctgtaacctt			
			cacacaggca			
_			ttttaattt			
			ataatcagco			
ttacttgctt	: taaaaaacct	cccacacct	cccctgaac	: cyaaacataa	a aacyaacyca	. 3300

attgttgttg ttaacttgtt tattgcagct tataatggtt acaaataaag caatagcatc 3960 acaaatttca caaataaagc attttttca ctgcattcta gttgtggttt gtccaaactc 4020 atcaatgtat cttatcatgt ctggatcgat ccccgggtac 4060

<210> 4
<211> 6116
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: R15/APPA
plasmid with pBLCAT3 vector

<400> 4 ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60 gagagteetg tttggtttaa geaacetetg ttteteataa acteeataaa aacaggaata 120 ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180 ataggtotaa taaccocgaa aatattacca tgatactgag catttgtaag tatctcatag 240 catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300 tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360 tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420 aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480 tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540 ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600 tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660 taaaacatat gtttgaccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720 gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780 gagteteaca aaatgaaaag gaaatatatt eagaaagaga atettgagag aatgtgttgt 840 cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900 taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960 ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020 tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080 gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140 atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200 tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260 gtaaaagaat aacatcatca ttcttaacaa tagaacatat aaagtcacac agaagtgaca 1320 gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380 tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440 cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500 gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560 attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620 aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680 agtatgaatt caaaggtgcc attettetge ttetagttat aaaggcagtg ettgettett 1740 ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800 agoogocaco atgaaagoca tottaatoco atttttatot ottotgatto ogttaacoco 1860 gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920 togtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980

cacataacca	acctggccgg	taaaactqqq	ttggctgaca	ccgcgcggtg	gtgagctaat 2	040
cacctatete	ggacattacc	aacqccaqcq	totggtaged	gacggattgc	tggcgaaaaa 2	100
agactaccca	cagtetagte	aggtcgcgat	tattgctgat	gtcgacgagc	gtacccgtaa 2	2160
aacagggaaa	accttcacca	ccaaactaac	acctgactgt	gcaataaccg	tacataccca 2	2220
ggcagatacg	tccagtcccg	atccgttatt	taatcctcta	aaaactggcg	tttgccaact 2	2280
ggetgatag	aacatgactg	acqcqatcct	cagcagggca	ggagggtcaa	ttgctgactt 2	2340
taccoccat	caacaaacaa	catttcqcqa	actggaacgg	gtgcttaatt	ttccgcaatc 2	2400
aaacttatac	cttaaacgtg	agaaacagga	cgaaagctgt	tcattaacgc	aggcattacc	2460
atcocasctc	aaggtgaggg	ccgacaatgt	ctcattaacc	ggtgcggtaa	gcctcgcatc :	2520
accyguacco	gagatatttc	tcctqcaaca	agcacaggga	atgccggagc	cggggtgggg	2580
aaggatgacg	gattcacacc	agtggaacac	cttgctaagt	ttgcataacg	cgcaatttta	2640
trtactacaa	cacacaccag	aggttqcccq	cagccgcgcc	accccgttat	tagatttgat	2700
caagacagcg	ttgacgcccc	atccaccgca	aaaacaggcg	tatggtgtga	cattacccac	2760
tteagtgetg	tttatcqccq	gacacgatac	taatctggca	aatctcggcg	gcgcactgga	2820
actcaactca	acacttecca	gtcagccgga	taacacgccg	ccaggtggtg	aactggtgtt	2880
tgaacgctgg	catcaactaa	qcqataacag	ccagtggatt	caggtttcgc	tggtcttcca	2940
gactttacag	cagatgcgtg	ataaaacgcc	gctgtcatta	aatacgccgc	ccggagaggt	3000
gaaactgacc	ctggcaggat	gtgaagagcg	aaatgcgcag	ggcatgtgtt	cgttggcagg	3060
ttttacgcaa	atcotoaato	aagcacgcat	acccgcttgc	agtttgtaag	gtataaggca	3120
attattaata	cccttaaacg	cctggtgcta	cgcctgaata	agtgataata	agcggatgaa	3180
tggcagaaat	togooggato	tttgtgaagg	aaccttactt	ctgtggtgtg	acataattgg	3240
acaaactacc	tacagagatt	taaaaaacct	cccacaccto	ccctgaacc	tgaaacataa	3300
aatgaatgca	attgttgttg	ttaacttgtt	tattgcagct	: tataatggtt	acaaataaag	3360
caatagcato	acaaatttca	caaataaagc	atttttca	ctgcattcta	gttgtggttt	3420
gtccaaacto	atcaatgtat	cttatcatgt	ctggatcgat	ccccgggtac	: cgagctcgaa	3480
ttcgtaatca	tggtcatago	tgtttcctgt	gtgaaattgt	tatecgetea	caattccaca	3540
caacatacga	a gccggaagca	taaagtgtaa	agcctggggt	gcctaatgag	tgagctaact	3600
cacattaatt	t gegttgeget	cactgcccgc	tttccagtc	g ggaaacctgt	: cgtgccagct	3660
gcattaatga	a atcggccaac	gegeggggag	aggcggttt	g cgtattggg	getetteege	3720
ttcctcgct	c actgactcg	: tgcgctcggt	: cgttcggctg	g cggcgagcgg	g tatcagetea	3780
ctcaaaggc	g gtaatacggt	: tatccacaga	a atcagggga	t aacgcagga	a agaacatgtg	3840
agcaaaagg	c cagcaaaag	g ccaggaaccg	g taaaaaggc	c gcgttgctg	g cgtttttcca	3900
taggeteeg	c cccctgac	g agcatcacaa	a aaatcgacg	c tcaagtcag	a ggtggcgaaa	3960
cccgacagg	a ctataaaga	t accaggcgtt	tcccctgg	a agetecete	g tgcgctctcc	4020
tgttccgac	c ctgccgctt	a ceggatacet	gtccgcctt	t ctcccttcg	g gaagcgtggc	4080
gctttctca	a tgctcacgc	t gtaggtatch	t cagtteggt	g taggtcgtt	c gctccaagct	4140
gggctgtgt	g cacgaaccc	c ccgttcagc	c cgaccgctg	c gccttatcc	g gtaactatcg	4200
tcttgagtc	c aacccggta	a gacacgact	t atcgccact	g gcagcagcc	a ctggtaacag	4260
gattagcag	a gcgaggtat	g taggcggtg	c tacagagtt	c ttgaagtgg	t ggcctaacta	4320
cggctacac	t agaaggaca	g tatttggta	t ctgcgctct	g ctgaagcca	g ttaccttcgg	4380
aaaaagagt	t ggtagctct	t gatccggca	a acaaaccac	c gctggtagc	g gtggttttt	4440
tgtttgcaa	g cagcagatt	a cgcgcagaa	a aaaaggatc	t caagaagat	c ctttgatctt	4500
ttctacggg	g tetgaeget	c agtggaacg	a aaactcacg	rt taagggatt	t tggtcatgag	4560
attatcaaa	a aggatette	a cctagatcc	t tttaaatta	ia aaatgaagt	t ttaaatcaat	4620
ctaaagtat	a tatgagtaa	a cttggtctg	a cagttacca	a tgcttaatc	a gtgaggcaco	4680
tatctcago	eg atctgtcta	t ttcgttcat	c catagitgo	c tgactccc	g tegtgtagat	4740
aactacgat	ta cgggaggg	t taccatctg	g cccagtgo	et gcaatgata	c cgcgagacco	4800
acgctcac	eg getecagat	t tatcagcaa	t aaaccagco	ca gccggaagg	g ccgagcgcag	4860

```
aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 4920
agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 4980
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 5040
agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 5100
tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 5160
tettactgtc atgccatccg taagatgett ttetgtgact ggtgagtact caaccaagtc 5220
attotgagaa tagtgtatgo ggogacogag ttgctottgo coggogtoaa tacgggataa 5280
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 5340
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 5400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 5460
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 5520
cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 5580
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 5640
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 5700
gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 5760
cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 5820
cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat 5880
tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata 5940
ccgcatcagg cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg 6000
ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttg 6060
ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg ccaagc
```

<210> 5 <211> 3470 <212> DNA <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: R15/APPA
. transgene

<400> 5

ggatccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60 gagagtcctg tttggttaa gcaacctctg tttctcataa actccataaa aacaaggaata 120 ctctttgtt ctagcataac caaaagatt agtgaattga aaacaatgtt cccttagagt 180 ataggtctaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240 catgtagtat ccatagtcca tcaatgagag agacatttaa catgatttc attaatcagg 300 tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360 tatttcacta aactaggttt atctatttg ttgctttctc taacatctct gcaatgaagc 420 aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480 tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540 tgggaagaaa ccatttggtg aacaatatt caaataaaaa tagacaacaa gaaaatatt 600 tgggaagaaa ccatttggtg aacaatatt caaataaaaa tagacaacaa tagttaat 720 gattctctt gggtggctgc aaattgtcca cgaataagac aaaatataaa aatgttaaa 720 gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgttgt 840 cacaaattaa agaaaacctg tggtgaatga catcctgagg ccttgagctat tactgacatt 900

taagataaag	gtaactgtat	acatttgtcc	cattgagggg	acaagaaagc	tgctctcatg	960
ttcagctcta	taattcttgc	cttaaacaac	ttaaatagaa	tgatttaaaa	tatggagctg	1020
tccatggacc	tttgaaatat	aaaatagtca	agcaacttat	caaggaatta	cagattcctt	1080
gatactaaca	caggtaaatc	ccacacgtgt	tttgagacta	catttgctgg	gattttattg	1140
atgtaatagg	tcacatgttt	ttcgggccaa	tgttgctgtt	attcggttac	ttcaagagaa	1200
tagtggcaac	tgatgctatg	tattctaggg	gtttgaagtg	atgtttcatg	attgaaattt	1260
gtaaaagaat	aacatcatca	ttcttaacaa	tagaacatat	aaagtcacac	agaagtgaca	1320
gtgtttaagc	tgtactattg	atcaaagaaa	tttattacct	tcagtttcaa	tggaaataat	1380
tactgataat	acaaacatgt	gtgaacacac	actaatccta	tccaaatgca	cagtgataca	1440
cagaaaatat	tagcaagtag	aatgcaatat	ttatataacg	attgtattta	tcaatcaatt	1500
gtatgtatca	atatatgggc	tattttctta	cacatgattt	tattcaaatt	tactctaatc	1560
attgttgaac	catttagaaa	aggcatactg	gcaacttttc	cttacctcat	ccagctgggc	1620
aaaagtccca	gtgtggagta	aaggatgcaa	gatttcctgc	tctgttaagt	ataaaataat	1680
agtatgaatt	caaaggtgcc	attcttctgc	ttctagttat	aaaggcagtg	cttgcttctt	1740
ccagcacaga	tctggatctc	gaggagcttg	gcgagatttt	caggagctaa	ggaagctaaa	1800
agccgccacc	atgaaagcca	tcttaatccc	atttttatct	cttctgattc	cgttaacccc	1860
gcaatctgca	ttcgctcaga	gtgagccgga	gctgaagctg	gaaagtgtgg	tgattgtcag	1920
tcgtcatggt	gtgcgtgctc	caaccaaggc	cacgcaactg	atgcaggatg	tcaccccaga	1980
					gtgagctaat	
cgcctatctc	ggacattacc	aacgccagcg	tctggtagcc	gacggattgc	tggcgaaaaa	2100
gggctgcccg	cagtctggtc	aggtcgcgat	tattgctgat	gtcgacgagc	gtacccgtaa	2160
aacaggcgaa	gccttcgccg	ccgggctggc	acctgactgt	gcaataaccg	tacataccca	2220
ggcagatacg	tccagtcccg	atccgttatt	taatcctcta	aaaactggcg	tttgccaact	2280
					ttgctgactt	
					ttccgcaatc	
					aggcattacc	
					gcctcgcatc	
					cggggtgggg	
					cgcaatttta	
					tagatttgat	
		_			cattacccac	
					gcgcactgga	
					aactggtgtt	
					tggtcttcca	
					ccggagaggt	
					cgttggcagg	
					gtataaggca	
	-				agcggatgaa	
					acataattgg	
					tgaaacataa	
					acaaataaag	
caatagcatc	acaaatttca	caaataaago	attttttca	ctgcattcta	gttgtggttt	3420
gtccaaactc	atcaatgtat	cttatcatgt	ctggatcgat	ccccgggtac	:	3470

<210> 6 <211> 5421

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: SV40/APPA +
 intron plasmid

<400> 6 cgagattttc aggagctaag gaagctaaaa gccgccacca tgaaagccat cttaatccca 60 tttttatete ttetgattee gttaaceeeg caatetgeat tegeteagag tgageeggag 120 ctgaagctgg aaagtgtggt gattgtcagt cgtcatggtg tgcgtgctcc aaccaaggcc 180 acgcaactga tgcaggatgt caccccagac gcatggccaa cctggccggt aaaactgggt 240 tggctgacac cgcgnggtgg tgagctaatc gcctatctcg gacattacca acgccagcgt 300 ctggtageeg aeggattget ggegaaaaag ggetgeeege agtetggtea ggtegegatt 360 attgctgatg tcgacgagcg tacccgtaaa acaggcgaag ccttcgccgc cgggctggca 420 cctgactgtg caataaccgt acatacccag gcagatacgt ccagtcccga tccgttattt 480 aateetetaa aaactggegt ttgecaactg gataacgega acgtgactga cgegateete 540 agcagggcag gagggtcaat tgctgacttt accgggcatc ggcaaacggc gtttcgcgaa 600 ctggaacggg tgcttaattt tccgcaatca aacttgtgcc ttaaacgtga gaaacaggac 660 gaaagctgtt cattaacgca ggcattacca tcggaactca aggtgagcgc cgacaatgtc 720 tcattaaccg gtgcggtaag cctcgcatca atgctgacgg agatatttct cctgcaacaa 780 gcacagggaa tgccggagcc ggggtgggga aggatcaccg attcacacca gtggaacacc 840 ttgctaagtt tgcataacgc gcaattttat ttgctacaac gcacgccaga ggttgcccgc 900 ageegegeca eccegttatt agatttgate aagaeagegt tgaegeecca ecaeegeaaa 960 aacaggegta tggtgtgaca ttacccactt cagtgetgtt tategeegga caegatacta 1020 atetggcaaa teteggegge geaetggage teaactggae getteeeggt cageeggata 1080 acacgccgcc aggtggtgaa ctggtgtttg aacgctggcg tcggctaagc gataacagcc 1140 agtggattca ggtttcgctg gtcttccaga ctttacagca gatgcgtgat aaaacgccgc 1200 tgtcattaaa tacgccgccc ggagaggtga aactgaccct ggcaggatgt gaagagcgaa 1260 atgcgcaggg catgtgttcg ttggcaggtt ttacgcaaat cgtgaatgaa gcacgcatac 1320 ccgcttgcag tttgtaaggc agttattggt gcccttaaac gcctggtgct acgcctgaat 1380 aagtgataat aagcggatga atggcagaaa ttcgccggat ctttgtgaag gaaccttact 1440 tctgtggtgt gacataattg gacaaactac ctacagagat ttaaagctct aaggtaaata 1500 taaaattttt aagtgtataa tgtgttaaac tactgattct aattgtttgt gtattttaga 1560 ttccaaccta tggaactgat gaatgggagc agtggtggaa tgcctttaat gaggaaaacc 1620 tgttttgctc agaagaaatg ccatctagtg atgatgaggc tactgctgac tctcaacatt 1680 ctactcctcc aaaaaagaag agaaaggtag aagaccccaa ggactttcct tcagaattgc 1740 taagtttttt gagtcatgct gtgtttagta atagaactct tgcttgcttt gctatttaca 1800 ccacaaagga aaaagctgca ctgctataca agaaaattat ggaaaaatat tctgtaacct 1860 ttataagtag gcataacagt tataatcata acatactgtt ttttcttact ccacacaggc 1920 atagagtgtc tgctattaat aactatgctc aaaaattgtg tacctttagc tttttaattt 1980 gtaaaggggt taataaggaa tatttgatgt atagtgcctt gactagagat cataatcagc 2040 cataccacat ttgtagaggt tttacttgct ttaaaaaaacc teecacacet eeeectgaac 2100 ctgaaacata aaatgaatgc aattgttgtt gttaacttgt ttattgcagc ttataatggt 2160 tacaaataaa gcaatagcat cacaaatttc acaaataaag cattttttc actgcattct 2220 agttgtggtt tgtccaaact catcaatgta tcttatcatg tctggatcga tccccgggta 2280 ccgagetega attegtaate atggteatag etgttteetg tgtgaaattg ttateegete 2340 acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 2400 gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 2460

tcqtqccaqc tqcattaatq aatcqgccaa cq	gcgcgggga gaggcggttt gcgtattggg 2520
	tgcgctcgg tcgttcggct gcggcgagcg 2580
-	tatccacag aatcagggga taacgcagga 2640
	ccaggaacc gtaaaaaggc cgcgttgctg 2700
	agcatcaca aaaatcgacg ctcaagtcag 2760
	accaggegt tteecetgg aageteete 2820
	ccggatacc tgtccgcctt tctcccttcg 2880
	gtaggtate teagtteggt gtaggtegtt 2940
	cogttcage cogacogetg cgccttatec 3000
_	gacacgact tatcgccact ggcagcagcc 3060
	taggcggtg ctacagagtt cttgaagtgg 3120
	tatttggta tctgcgctct gctgaagcca 3180
	gatccggca aacaaaccac cgctggtagc 3240
-	cgcgcagaa aaaaaggatc tcaagaagat 3300
	agtggaacg aaaactcacg ttaagggatt 3360
	cctagatcc ttttaaatta aaaatgaagt 3420
	cttggtctg acagttacca atgcttaatc 3480
	ttcgttcat ccatagttgc ctgactcccc 3540
	taccatetg geoccagtge tgeaatgata 3600
	tatcagcaa taaaccagcc agccggaagg 3660
	ecgecteca tecagtetat taattgttgc 3720
	latagtttgc gcaacgttgt tgccattgct 3780
	gtatggett catteagete eggtteecaa 3840
	tgtgcaaaa aagcggttag ctccttcggt 3900
	gragtgttat cactratggt tatggragca 3960
	staagatgct tttctgtgac tggtgagtac 4020
	eggegacega gttgetettg ceeggegtea 4080
	actttaaaag tgctcatcat tggaaaacgt 4140
•	ccgctgttga gatccagttc gatgtaaccc 4200
	ttactttca ccagcgtttc tgggtgagca 4260
	ggaataaggg cgacacggaa atgttgaata 4320
	agcatttate agggttattg tetcatgage 4380
	aaacaaatag gggttccgcg cacatttccc 4440
	attattatca tgacattaac ctataaaaat 4500
aggegtatea egaggeeett tegtetegeg o	egttteggtg atgaeggtga aaacetetga 4560
	tgtctgtaag cggatgccgg gagcagacaa 4620
gcccgtcagg gcgcgtcagc gggtgttggc g	gggtgtcggg gctggcttaa ctatgcggca 4680
tcagagcaga ttgtactgag agtgcaccat a	atgeggtgtg aaatacegea cagatgegta 4740
aggagaaaat accgcatcag gcgccattcg	ccattcaggc tgcgcaactg ttgggaaggg 4800
cgatcggtgc gggcctcttc gctattacgc (cagetggega aagggggatg tgetgeaagg 4860
	cagtcacgac gttgtaaaac gacggccagt 4920
	gtatgttgtg tggaattgtg agcggataac 4980
	attacgaatt cggcgcagca ccatggcctg 5040
	gtaccttctg aggcggaaag aaccagctgt 5100
	ccccaggctc cccagcaggc agaagtatgc 5160
	ggtgtggaaa gtccccaggc tccccagcag 5220
	agtcagcaac catagtcccg cccctaactc 5280
	ccgcccattc tccgccccat ggctgactaa 5340

ttttttttat ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt 5400 gaggaggctc gaggagcttg g 5421

<210> 7
<211> 17732
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Lama2/APPA
transgene

<400> 7 tegagagtat etttgteage tgtgeeteea acaaaggggt actgttgeee acatagaaag 60 atctaaacta attaattaat ccctcacccg caaatctttc agtcactaag ttagcacgat 120 tgttgaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180 ctgaggagac acctgcatct gactaagaag agccacggtg ttagttgaat ggtgtggagt 240 agggtggttc tgtgggacag tagaaaatcg agaggcatgt gccgtttagt gaactgatgg 300 aagctacccc aaacgacaga gattgtcagt caggccaatc cgtttcgagt ttgatgggca 360 gccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420 agggattgag agaccetgae aggegeaagg ceetaacac cacacetace aceteaettg 480 acaaagetge caaagaccaa agaettgtte teeattagaa atgacagetg gettgaceeg 540 acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgtttc attcagtatt 600 ataaaaggac agtattacag attttgttgt acactgctgt tacatgtggg gcagtgtgtc 660 tttaagtagg gtaaagtact ctttaaaaat gggtcctaga tattttttcc tttaactcaa 720 gtotottact gtttaaatga tttttatttt gtttaatatg gaggaaaaag aagogtaaat 780 ggacaatata tatttagaga aagatggtta gctgtcagaa aaatatgcaa atcaaaatca 840 caccaagact gcagcacacc cctgtcagat ggctgtgatc aagaaaataa atgacaatga 900 cacactggag caaccactgt ggaaatcagt atgaatggtc ctcaaaaacc tgaagataga 1020 geggggegtg gtggcataca cttttattcc cagcactggg gaggcagagg caggtggatc 1080 tetgagttee aggecageet ggtetatage acaggtteta ggacagecag ggetacacag 1140 accaaaccaa accaaaccag accaaaccaa aacactgaag atagaacttc agtattccat 1260 tectagatat atacceaatg gagactaagt cagcaagaca eetgeacage catgtteact 1320 actacactgt teaceacage caggetgtgg aaccageetg agtgteeatg ataaatgaat 1380 ggataggtaa ctttcaaggt aaatggactc tgctgtgtac atgcctcaca ttctgtttat 1440 tcatttttct ttatgaggtg tccattcagg agtcacatgg tagttctatt ttcagtcttc 1500 tgaagatact acactggtcc ccacagttta cacttttatc agcagtgaat aagggttcct 1560 ctatccttac catcatttgt tgtaattttt cttgatgacc ctctttctga cagggatagg 1620 atgtaatatc agtgtgagga agtacaactt gttttctaag tatttattgg ccccttgcat 1680 ttettetttt gaaaactgte ggtteetgae atetgeteag gtatteattg gatgttgttt 1740 ctttggtgtt tgagttctta tgaattctag atgttaaatc cctgcctgtg gttctctccc 1800 attetgtagg etgeeteete accetggeaa ttgttgteet tgttttgeag aaaettttga 1860 cttcatggaa tctcatttgt cagttttccc tcctctgcta tagcctgagc taatgcactg 1920 gtttttacag agccctggtc tatgccttta tcctcctctg gcagcttcgg agtttcattt 1980 cttacattta gatctttgat ccactttgaa caagttttgg agcagggtga gagatacgaa 2040

tctagttcca	ttcttccata	tgtgatccta	gtttacatag	catcgttggt	tgaagaggtt 2	100
ttattttatt	tttaaataat	gtgtcataaa	aaacgaggtg	gttgtagcag	tgtggatttg 2	2160
tttctttqtc	ctttgatcta	caggtcttgt	tttgtgtcag	tctcatgatg	ttttattgct 2	2220
atogctctqt	catacagtct	gaggtcaggt	attgtgatat	accttcagta	ttgctccctc 2	2280
agactcaggt	ttqctttggc	caggagtcat	cttactcagt	gctcttagag	ctcccccagc :	2340
atgtagetge	tactattctt	agttgataaa	tcaggaaact	ggggctcaga	gagattaact :	2400
gtcttgaact	acttctgggg	aggtgaaacg	tggagacact	aaactgtgtt	taccctgtac :	2460
toctccaqta	qctqtcgggt	gctgggctac	agcaaagcac	ctatactata	tattactcag	2520
gaggtggaaa	aactcagcct	cccttggggt	tcccaagctc	ccaggtgtcc	agtcactgct	2580
ggaaacctca	tggagtctga	aaggaagggt	tgagggtaca	tggggcagcg	atgaggagcc :	2640
tggggctggg	atctcccaaa	cacctggata	tccagatgcc	actgggtcag	ggggagttgg	2700
gaacagagtt	gggatgtcca	tggacctgtg	acaaggccag	ggccaggggg	aggataactc	2760
tggctttact	aatttgcgaa	agtccttagc	ttagcagcag	ttgtctggga	gcacagaggg	2820
gccttctgta	agaggeteag	gcagtgccgc	tctgtaggcg	aaggtettet	ccatgttccc	2880
catggtggtt	cttgatgaaa	gagacagtcc	ttggctccaa	actggtttat	tgattgttca	2940
ttgtggaaaa	tgggtgcaca	ccaccttctc	agggtggacc	agagatcaaa	taccttttgc	3000
agggaggaat	atctgggaag	ggacgcttac	tggctaaacc	ctcagggcct	ctagatacat	3060
cattagcatg	gagaactctg	ttctgggcta	catgaccaca	ggccacattt	ccacaagcca	3120
catgtgggaa	gtgtggcaca	tgttctaggc	caggaatctg	gtagggagcg	tggagccacc	3180
taccatccca	ggtgggtgcc	tgggtgccag	ggaccctgaa	cccgctcaac	cttaccaagt	3240
ttcctggcag	ggtccactgt	cctacacaga	agctggagga	ggtgtgaggg	ttgtgtcttt	3300
gtggaatgtc	ccatgctgct	tggggctcag	tttctccacc	tgtacctcat	tggtttgggt	3360
ataaaaagtg	gggatacttt	attattctct	gactcggtcc	tgaggaaaaa	gcatcgtggc	3420
agtccaggaa	ccacaccctg	aggttcctgc	actgaaggga	ctccctaagt	ctctggagtc	3480
teteceette	acagagetge	caaagtctag	gttcttttga	ggataacaga	gccatgcttg	3540
gtaagcagac	aacagcattt	gtttactcaa	ccttcttttg	tcagctccct	cttcataaac	3600
aagttgagad	accatgctgg	cttgaggaag	acttctaaag	ccagacaact	gtgcaaggaa	3660
gaagaagaag	gggcaagtgg	agttagcctg	gatgtagccc	tcaaagtctc	cagagaccag	3720
ccatgaaggo	: tcaagtggag	ggcaagacct	geageageca	ageatetgge	aggagaggat	3760
cctgggaacc	cctctaccat	gacacacatt	cttcctgcag	gteacaetta	ataggccatt	3040
tettatttgg	g atctatcatg	gtgttctgtg	g cgagattaat	gaggtgttat	gctgcgaaca	3960
gaaagttata	taaaaacaag	tececece	ttgtcactgo	e tgctaagaal	gtagcagaaa	4020
ttgtctcaag	g tgtctctcta	atcagaaaca	ataaaggtct	cettggatte	aagccctcca	4020
gtttcctcct	tccttgctga	geettggae	a cccatacaaa	a ceteetyyat	gctacagctc	4140
tgggcagaga	a ctccaaggtg	gggagagaci	t gatggtacae	tectecaea	cttgtttggg	4200
ggtacaccca	a ctcctctgcc	egrgrggra	c etgeagleag	, coordatace	aggeeeteag	4260
tgggtette	e atgggeaaca	cgcagaggg	a gycaacygai	t etaceetet	acaccctggt	4320
tagtttacce	e eggeeatget	. etetgetet	c cacecetet	t ctcctcacc	g ccacggettt c tagetatgat	4380
ctctgcagg	a attacatett	. catattyyt	a ataaasta	a ottotagaa	g gctcacggag	4440
gtttaettt	a gagtgacett	totaastas	t gegggddeg.	a agttectog	t ggttgtctct	4500
acgetaggg	a agaaacycci	, cocadolado , acottosao	t tagtgeace	a gaaatcaga	g cccagtcaca	4560
geetteee	t cycladayto	. accurgady	a aatacceea	a atccatgac	a tttcaagagc	4620
gaytaaata	a cottagete	tttaaaatt	g tatettete	c ttgatgtat	g tgtgtgtgtt	4680
tatassar	a transfort	, cocyyyycc r acataanan	a tactetasa	g gacagagag	g atttgcaatt	4740
caccaaaga	g egagatyyti	acettacte	t agtaccaga	a qqaactqat	g cagaaaagag	4800
graycatgt	y acatected	g goodatagood	a tagaggaga	t cttacagto	c attccctcct	4860
caayayytt	a colouryday	a gergeeder	t gactgcaac	t gaaacagag	g cttgggatgg	4920
ccaggccct	g cergaggar	a yacacycyc	.c gacegeaac	- 55-243	J = JJJ - 755	

•				.		020
agagttaggt	tcacagaagg	gagggtggga	gatggatget	rgergggtte	tgggtctcat 4	-040
caccagctcc	tgaccacccg	gtcagcccat	gtgcttattc	catagette	ttttgctatg 5	-100
tttactcagt	gtggtgtttg	ttgggaccca	gcagaagcca	gtcccaggct	gacagetgtg 5	2100
gatacacagg	gcagcatgag	ggtcctcagc	ctgaagcagt	caggctggca	gaagagaaag S	2100
accagcacac	attccttcaa	ccaactatgt	cttgaaaaac	aaacatatta	tatcacatat !	5220
attgcattta	tgagacagct	aaaatgtact	cgggtagcat	gactccaggt	ggggatatct !	5280
gcaagtgcca	tgagtggcag	agggacagcc	aatgtgaggc	aagaaggaat	tctggctcaa !	5340
cacagettag	ctccctggtg	ttggttcaaa	ctttgagagt	ttgaccacaa	gcactttatt !	5400
tttgacatat	ttaaacagag	cacaactttg	ggaaaaagtt	ttcttatgaa	aattatcaca	5460
ataaagctta	aggcatgact	acattaaaat	gcctttgcaa	agtatatgtg	ccctcttcca	5520
caagaatggt	tctattgact	gagaaataat	gttcaggata	aagatccagg	aagaaaagat	5580
cagggataag	taaaatacta	aactcttttg	caaagtacat	agaccctctt	tcataacaat	5640
gggttctatt	gactgacaag	cactgctcag	gagttgggaa	agagtctagc	ataagcacga	5700
tagcctggag	actctagtga	ggtctagtct	tacagacagc	aaaaatcacc	aggttacaaa	5760
ctacattcat	ttccagtttt	ctgatcaggc	acaggtatga	atcccttctg	ttgaagagaa	5820
aagtccatgt	gtttaaaata	tctggtttct	ccagtgctat	tagcgagaag	acttgagccc	5880
tatacaacto	ccacctggag	tgacatcctg	tcttcatggt	atattacata	cctagacacg	5940
ctcatctcac	agacttagga	ctttgtcttc	tgatctccat	ttctgatccc	acttccacct	6000
ttgccttgat	agtgtcattt	tcttcactgc	cttggtgaca	accatgttat	cctctgtgta	6060
tttgagtgtt	accattttca	gattttacct	gtatgcaaga	tcacacagtc	tttgtctttc	6120
tatctagate	catgctaato	tctacacaac	aacccttccc	cgtcactcag	atcttcctcc	6180
attaacacat	acatggtgct	gaagaggcta	gggagcttcc	cttcagtggg	gagctagctg	6240
getattggg	ctttttgact	gtccaggaag	gcccccaatt	gctgagacaa	gaacttagat	6300
tetteattat	tgactctaac	tcatgtatca	agcagaagct	aatgaatagt	tatcaacagg	6360
atcagaggtt	t ccagtgtaag	acactttgac	atgaaagaac	: ggaggaagga	cagatggatg	6420
cataaaagca	a ggaccactgo	cccaggaagg	tcctggaaac	: tgatgcaggg	caaaggacag	6480
gttataaac	c aaatcttagg	gagtcaggaa	gagcacagag	gagctcaacc	aactgaccac	6540
tgcttaggg	g ctaccaacco	aatcctccct	gtgggaacag	g ctaagctato	agccaagggt	6600
aataaacag	g caggacctgt	ggatgacatg	gagagcatag	ggaccctggg	tccagccttt	6660
agcacctgc	a ctctcaggat	actccaccat	tgtgtcttag	g agageetage	gatactgggt	6720
ccagccttt	g gtaccttca	tctcagggta	a ccccatcact	t gtgtcttgga	a gagcctaggc	6780
accetgggt	c cagcetteag	g tacctgcgct	t ctcaggacae	c cccaccattg	g tetettgeee	6840
cgtctcttc	t teetettee	t ccctttcatt	gtetettete	c tgtttcttt	ttgactctcc	6900
					c ctattctctc	
tgtggtccc	t ccactttcc	t ttatctctc	a tgcttctct	c ctccctcaa	a tacttgtcac	7020
ccactatac	t tcaggggcc	a gctctagtg	a caaagctgt	t aatagcaaga	a ctctcagatc	7080
tccaacggc	t cagaggagc	c agacccacc	a agaactctc	t ccaggtcca	a tttcaggttc	7140
cttcgaaag	c tttcagcaa	a tgctcaggg	a acatgccac	t aacaagaag	a tgcaaattcc	7200
agttgagag	t gggaaaggc	c cttgcgtag	g teceatett	c caggccaag	g tcagaggggc	7260.
tetqtqtaa	t ceggattga	c agggeteag	a acaatgttt	t gtttttaag	g tttatttatt	7320
ttaggtgtt	a gtgtctttg	c ttgcatgac	c ttatgtgca	t catgtgtgt	g caggttcctg	7380
atgacagta	g aggagggct	t tgaatccct	g gggatagga	a gttacagga	a attataagct	7440
qctttatac	g tettetage	t ttcccaaca	g aagtgaatg	c tcttcacca	c tgagccatct	7500
ctctaggc	c aagagacat	t gctttatgg	a tataattgt	g tgtgtgtgt	c aacattgagg	7560
aaagggaaa	nt aaaaaaaaa	a cttcagccg	c taaggttgt	a cagtttcac	t aattgctact	7620
tttagttgt	g ataaaatg	c aggtgcttc	a acatttata	it atacaaaaa	c ttccctgctg	7680
gtgattca	ac totoagaac	t ggggtaagt	g ggtgagtto	t cttttctg	t ctctgtctct	7740
gtetetet	c ttccattct	t tettaaagg	a aataaacat	t gcagctggg	t tatageteat	7800
2			•			

	_ 	.		ataastaata	ttaggaggta	7060
				gtgggtggtg		
				tctgtctctg		
				ggcctggttc		
_				ctgagtggtg		
				aggaagctgc		
				cagctaatcc		
				tggccccatc		
				acttgatgga		
				ctgacagcct		
caacaagggc	cctctatgtt	tgctatgtaa	tgtaatgtca	gacattgtca	ggagtgtccg	8400
cagcacagcc	tgcccagtgt	gagggctctc	ataggtttcc	cactgtctta	tctacacagg	8460
gataacgagg	aggtaagctg	cagttcccag	tctcacttca	cagaggaaga	gataacccca	8520
tcccaggtca	tgtagccagc	agtggaaaga	atgaggattt	gaactcaggt	cttccaagtc	8580
ccattgatag	catctcctca	caagtccctt	gccaccctca	cgatgcctta	gacacttgcc	8640
tgccctttat	actaaggaga	tgcaggtaca	aggggtttac	ccatġtagca	gctgaggcag	8700
ctggggatag	ataccagcag	caggcctgat	gtcaccactc	taactccagc	atccccagtc	8760
tgtgttcctg	gagtgtgaaa	atccctactt	aacaagattg	tgcaacagtc	cttggctctg	8820
tgacccatag	ctggaaacag	gattctcatt	gatttgtgga	acatggtggc	agccagccaa	8880
aaagagggtc	tgcatacaga	agacacgtgt	ggcaaggcca	cagcagactc	tgactacctt	8940
agcttacaga	attacaaggt	cataatgtcc	tctgctttgg	tcacctcatg	ttaaggacag	9000
gccctaatga	agatggggca	gaagactgaa	ggaatggcca	accaataact	ggcccaactt	9060
gagacccatc	ctacaggcaa	gcatcaattc	ctgacactac	taatgatact	ctgttatgct	9120
tgcagacaga	agcctagcat	aactatcctc	cgagaggtcc	acccagcaac	tgactgaaac	9180
agaaaaagat	atccacaggc	aaacagtgga	tggaggtcag	ggactattat	gggagagctg	9240
tgggaaggat	taaaaaccct	gaaggggata	ggaaccccac	aggaagacca	acagagtcaa	9300
ctaagagacc	tgtgggagct	ctcagagact	gagccaccaa	ccaaagagca	tacacaggcc	9360
ggtccgaggc	acctggcacg	tgtgaagcag	acatgcagct	cagtctccat	gtaggtcctc	9420
caataagcgg	tagcctgact	gcagtatcca	atccccaaca	gggctgcata	gtctggcctc	9480
agtgggggag	gatgccccta	atcctgcaga	gacttgatga	gtggagagct	atccaggggg	9540
aacccaccct	ctctgagaag	ggaatgggga	tgggggaggg	actctgtgaa	gaggggacaa	9600
ggacaaacaa	gaacctcaaa	taggtcaggc	cctaaaggct	tgctaagtag	cagtggccca	9660
gctctgtcct	gttcctcagc	ccaaggctca	gctcccacct	gtttctgtgt	ttttctggct	9720
tttcatgggc	ctaggacttg	gtgaccagtt	caaacaatgg	ggcctgtgga	agacacaata	9780
tacaagacta	gggacattcc	tgttctgctg	actatccata	gcctgatgta	ggtggaagga	9840
cccaatcact	ggatttctac	ccttgcacaa	ccttgacago	tgagggcctc	tcagaaacct	9900
atttcttcca	ctgaaaaatg	agactctcaa	atgaacgtcg	, tgacaatcat	caggettatt	9960
aaagaggtgt	atctaacctg	aatggcaagc	agacagcagg	g caaatgtctg	tatcaacctc	10020
taggaaggad	: aagaactgct	cactgctgcc	cccaggagg	, ccatttgctg	aaacagctgc	10080
teteetgete	gtgcacaggc	cctgccttct	cattgcagco	acagcccctt	cctgtctgaa	10140
cctcctgtca	ggtcactggg	aaacagatca	agatggaaca	ggacagetee	: tgatggtaaa	10200
taaaaaacag	, tggtcatggc	tattcatagg	ggtttatgct	tetteagtee	acactgtgaa	10260
gagctgtggg	catgaaccac	agtgttcgag	gtagagttgg	g ggttctgaaa	ttcacagtgg	10320
				a cacacagtco		
				agaggccaca		
				g tgtgtgtgtg		
				t acacgttgag		
				c tggagaacaa		
				t gctcacatac		
•						

gacaatggag	cctccatcag	agtattactt	tagctcctca	ccgatggcaa	tgcaccacct	10740
	ataggagttg					
	tatctggtaa					
	cctggatcta					
	ttcccccta					
	ccacagacat					
	tggctcacag					
	tccccagtc					
	tttcccctcc					
	gggcagcatt					
	gaaagagtga					
	ggttgaagct					
	cacaattgaa					
	ctggctagct					
	tctggatctt					
caagtttgac	aacatagggc	tttgaacttg	gcacaaggtc	catcactgtc	acccaagcat	11640
cctgggtgac	ctttgggttg	gaatatcttg	gctaacctta	gatattttct	ttggagtatc	11700
	ccaggaaata					
	ggagatcgtg					
	agtgactcca					
	cagagggaag					
	, atgacattgt					
	tttattcat					
	tcaataatga					
	g gagaaaggca					
	a tagaaggtgt					
	a ggcaaggcag					
	acacacattt					
	gctaagcaca					
	a ttgtggctga					
	g gcatatcccc					
	a aggcaagcgt					
	a ctctccattt					
	c ccatttttat					
gagtgagee	g gagctgaagc	tggaaagtgt	tatasaccas	agregetary	, gegegegege	12840
tccaaccaa	g gccacgcaac	caracaga	taataaasta	ategectate	tcggacatta	12900
ggtaaaact	g ggttggctga g cgtctggtag	cacegegegg	actaacaaa	accycctac	cocagteta	12960
	g egtetggtag g attattgctg					
ccaggeege	g gcacctgact	atgregates.	catacatac	caggcagata	a catccaatc	13080
egeeggget	a tttaatccto	taaaaactoo	. cgtttacae	e ctggataac	g cgaacgtga	13140
tasaaast	c ctcagcaggg	. cadaadaata	: aattoctoa	tttaccaaa	c atcqqcaaa	13200
	c gaactggaac					
	g gacgaaagct					
rarrara	t gtctcatta	ccatacaat	: aageetege	a tcaatqctq	a cggaqatati	t 13380
tetestas	a caagcacag	asstaccan	a gccagaata	g ggaaggatc	a ccgattcac	a 13440
ccantanaa	a caagcacag	gtttacata:	a cgcgcaatt	t tatttocta	c aacgcacgc	c 13500
	c cgcagccgc					
agaggeege	.c cycayccyc;	,		J, J 		

oggacacgat actaatctgg caaatctgg cggccagtgg tgaactggtg tgaactggtg tgaactggtg tjaactggtg tjaactggtg tjaactggtg tjaactggtg tjaataa ggccgtggg 13740 aagcgataac agccatggga ttcaagttte cagactteta cagactteta cagactteta cagactteta cagactteta cagactteta cagactteta taatcggaag ggacaccacca tectaccacca tectaccacca tectaccacca tectaccacca tectaccacca agaccaccac agaccaccac agaccaccac acatteta tectetetag tagatteca tectacctga 14100 acacaccaca agaccaccac caattetact tectetetag tagattecac tectacacca tectaccacca 14100 acacaccaca agacaccatc caattetaca tectacacca acactteta 14100 acacaccaca agacaccatc caagcattga tectacacac caattetaca tectacacca caccttetag tetacacaca tectacactag tectacactag tectacacaca dacattegag tettacactag tettacacaca acacttetag tettacacaca tettacacacag t	ccatccaccg	caaaaacagg	cgtatggtgt	gacattaccc	acttcagtgc	tgtttatcgc	13620
oggetagege gataacacge ogccaggtgs tyaactggtte tytaactggtte gegetaget 17474 aagogataaca agccagtggat tyaactggtte gegetagate 13800 tyaaaacacg cogetyteat taaatacgec geggaatgg gyaactaca cetygaagg 13800 tyaagagaag cyaaatacge agggcatgt teetytegea gytttace aaateggaa 13920 tyaagagaaga cacacacact teetecty gyatacacac teetecty 1400 aagagagaaga gecacacacac teetecty teetecty 14100 gaccaccaca agaccaceca agaccaceca teetecty 14100 gaccaccaca agaccaceca gatetteca teetety 14160 gaccaccaca gagcagty coaggegyte tyaactecaca cacety 14160 gaccaccaca gecagty coaggegyte tyaactecaca catety 14220 aagagaaga dytyggatte teetyggatte gyactecaca cettyggata teetyggata teetyggygtte gyactecaca teetygaaga ty							
agogataaa agcagtgga ttaggttte getggtette cagactttac agcagtgcg 13800 tgataaaaacg ccgetgteat taataacgce gecogagaag ggaaaactga cectggcaga 13800 atggaagaag cgaaatgcge aggactgtg ttegttgga ggtttacge aatceggaa 13920 tgaagcacca ataccogett gagtttgta aggtacccgg ggatcacaac ttecacctg 13980 aagaggaaga acagaaggat gccacaacte tectgetgge tactectcaag tggtttcate 14040 ttacttetga tggcatttee cettagaaag tgctacctac accacacaca tectacccaa aggaccacca caaattetet tectetetga gtagtetcac cacetgtac 14160 caccatecca gaattaaaat ectaactga cettgggtg tgacttgac cacetgtge 14220 aataaagagt gttggcagtg caagetggg tgggggatg tgactgaagggagaga ggaaaccag etttaattca agcacttggg 14280 aggacagagga ggtatacaaga gaaaccctgt gtgcgaaaac caaaaaaaa aaaaaagtt 14400 gttggcagag tgtgggtat ataccaggtg ggaatttcaa atgagtggg tagggggatt tcagggggggggg							
tgataaaacg cegetgteat taaatacgee geceggagag gtgaaactga cectggeagg 13860 atgtgaagag cgaaategee aggegatgtg ttegttggea gyttttacge aaategtgaa 13920 tgaagcacge ataccegett geagtttgta aggtacceg gytttacege aaategtgaa 13920 tgaagcacge ataccegett geagtttgta aggtacceg gydtecacac ttegectett 13980 aagaggagaga acagaaaggat gecacaacact teetsgetgge tactetecag tggttteate 14040 ttacttetga tggeatttee etctagaaag tgetactate accacacat teetacetga 14100 gaccaccaa aggaccece caaatteett teetsgetgge tacteteca cacctgttae 14100 gaccaccaa aggaccace caaatteett teetsgetgge tgaettgeet cagteettge 14220 aataaaggt gttgggagtg cagaggtgg tggggacge cettaaattge aggacgaggg gageggagg geggagtg geggagagg geggagggggggggg							
tgaggaagag cgaaatgeg agggatgtg ttegttggca ggttttacgc aaatcgtgaa 13920 tgaagcacgc atacccgett gcagtttgta aggtaccegg gatcacaac ttegcettga 13980 aagaggaaga acagaaggat gccacaactc tcctgctgg tactcatca ttetacctga 14100 gaccacccaa aggacctcc catactgca ttetgctgt gatttcca ttetacctga 14100 gaccacccaa aggacctcc caattctct tcctctctga gagttcca cactgttac 14160 caccatccca gaattaaaat cctaactgac tcctgcggtg tgacttgcc cagtccttgg 14220 aataagagtg gttggcagtg ccaggcgg tgggcacgc ctttaatcc agcacttggg 14200 aggcaaggc aggcggattt ctgagttcga ggccacgcctg gtctacagag tgagttccag 14340 gacagccagg gctatacaga gaaccctgt gtcgaaaaac caaaaaaaa aaaaaagtt 14400 gttggcagag tgtgggtat ataccaggtg gagattcaa atgagtggct gagggcaggacgaggagacgaggaga cttagaggat agctcataac ttaaaaagaa atgagaggac agacgagaga 14520 attggagaga tgggcacaca gccactgtg gaatttcaa atgagtggct gagcgcacgc attgagagag tgggcacaca gccactgtg gaatgtgga gaaccaaca cagcacatga 14500 cacagaaggaa ctgggcacaca gccactgtgt gaatgtgga gaacacaac cagcacgat 14580 acatgcata agtgtatatt ggcgccatcc tgactgatga gacacaggaa aacagatag 14640 cggggttagg tgggcacagg ctttcctgc tgctcttcc taaggtcat ctcaagagct 14760 ccttgggtac atcaagaggt cttctctagc tgcctcttcc taaggtcat ctcaagagct 14760 ccttgggtac atcagtgatc ctggtgatat ccagggcttc ctgattccat ctttgtcat 14760 ccttgggtac acagtgatg ggaactcca aggacataca ggagaagaa caagaagaagaagaagaagaagaagaagaagaagaagaag							
tgaagcacgc ataccegett geagtttgta aggtacacaca tectetceag tggttteat 14040 taagtgagaag acagaaggat gecacaactc tectetceag tettetceag tggttteat 14040 taactetcta tggcattecc cetatagaag tgetactaca tetctacctag 14100 gaccacccaa aggacctecc caagtette tetctettetag tgagtteca caccetty 14100 aataagagt gattagaat cetaactgca ctetagegtg tgagtteca cacettyte 14220 aataagagag gagttaggt ceaggegtgt tggcgaccg ttetattec agcacttygg 14280 gacacgagg gettggggat ceaggggttgg tgcgaaaaa caaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 14400 gtgggaag tettagagga gagttacaaaa gaactgagga tgcacatgg taacaagaga taacaagaga taacaagaac tacaagaac tacaagaac tacaagaac tacaagaaca tacaagaaca tacaagaaca tacaagaaca tacaagaaca tataaaaaaaaaa tacaaagaaga tattacacaa gaacac	-						
aagaggagaa acagaaggat gccacaactc tcctgtgtgc tactectcag tggtttcact 1400 btacttctgta tcggatttcc cctctagaaag tgctactact atccacacca tctactgta 14100 gaccaccaca aggattcc caattctct tccttggegtg tgacttctcc caccgttca 14120 aataagagtt gttggcagtg ccaggcgtgg tggccaccc ctttaattcc agcacttgg 14230 aggagagaga aggcggattt ctgagttcga gccaagcctg gtttacagag tgagttccag 14340 gttggagag tgtgggatta atcacagggg agaacccatg gaaaccagaa aaaaaaaaaa aaaaaaaaaaa 14500 gttggagag tgtgggatta agcacatgag gagatttcaa agcactatac taaaaaaaaa tagagagaa 14500 ccagaaggag tgtggacaca gacactatac taacaagaga aacacagaac tagcagaaac tagcagaaac tagcagaaac tagcagaaac 14760 cagggttaag tggcattcatac tatgctactt agcattcataa agcattcataa agaagatta 14640 <							
ttacttctga tggcatttcc ctctagaaag tgctactacca tcctactcga 14100 gaccaccaca aggaccatcc caaattctct tcctctctga gtagtctcc cacctgttac 14160 caccatccca gattacaaaa cctaactgca ctctgggtgt tgagtctcgcc ctttacttcc agcacttgg 14220 aataagagtt gttggcagtg ctgggtcacc ctttactacc agcacttgg 14300 gacagcagg ggtggttac ctgagttcga ggcaaaaac caaaaaaaaa aaaaaaagtt 14400 ccagaaggaa cttgaggtaa acccatgagt gaaattccaa aaaaaaaaa 14400 ccagaaggaa ttgaggata acccatgag gaaattgaga tagactgaa 14520 attgagaag tggccatac gaattgaga gaacaaaga aacatgaga 1460 cagggttaga tggccatacc tgacttgaaa gaacacaaga aaaaagattga 1460 cgggttaga tggcacatcc decttcgcc tgccttctga gaacacaaga aaaaagattga 1460 cgggttatect tacatettac acttcatac							
gaccacccaa aggaccctcc caaattcatct toctctctcta gtagtctcca caccttctc 14160 caccatccca gaattaaaat cctaactgca toctaactgca cagtccttg 14280 aataagagtc gtggcagt ccaagcgtg tgggcaacc cttaattcc agtccttgg 14280 aggcagagg aggcagagc gaatctcaga gaacccatg gtctacaaaa taattacag 1340 gtaggaag gtdgggttaa ataccaggtg gaatttcaaaa ataaaaaaa aaaaaaaaa 1460 ccagagaga cttagaggat actcaaggta gaccactgtg gaattgagga gaaccagaaac 14500 attgagaga tgggcatagc ceactgtgt gaattgaga gaaccagaac taccaagaac gaccactgtg gaattgaga gaaccagaaa 14520 attgagagag tggccatagc cttctctgc tgcctcttcc taacgagaga cacaagact 14700 tatggttag tggccatagc cttctctgc tgcctcttc taatgaggac accaagagaa 1460 ccttgggtta taacatgtga cttctacacc agattcata							
caccatecea gaattaaaat eetaactgaa etetageggg tgaettgeet eagetettge 14220 aattaagagtt gttggeaggg ceaggegggtgtg tggegeaege etetaattee ageaettgg 14340 gagagagge aggeggatt etaagtga gecagectg gtetacagag tgagtteeag 14340 gttggeagag getatacaga gaaaceetgg gtegaaaaac caaaaaaaaa aaaaaagtt 14400 gttggeagag tgggggtta ataceaggtg gagattteaa atgagtgget gaagetgaag ettagaggag ettagaggat ageteataac ttaaaaagaa atgtagaggg tageagaaac 14520 attgagagag tggggacaca gecaetgtgt gaagtggag gaacacaate eagecageta 14500 tacagaaggaa tgggacacaa gecaetgtgt gaagtggag gaacacaate eagecageta 14500 tacagaaggaa tgggacacaa gecaetgtgt gaagtggag gaacacaate eagecageta 14500 tacagaaggaa tgggacacaa gecaetgtgt gaagtggag gaacacaate eagecageta 14700 tatgeteet taactettee attgetaett agettetaga tacacacee agaagaaga 14640 ecettgggtaa atcagtgate ettgetaett agettetaga tacaceete agattagte 14700 ecettgggtaa atcagtgate ettgetaett ettgetaeta ettgggataa agaagagtegaagagagggtgaagagggtgaaa cettetaata ettgacacee etaagggetea agaagacae 14800 agaagtteac agaggtgaag tgatteatga gagacatacaa gtgagacaaga atcaagggte 14940 ggattatetg actetaetet aacttttatg taaatgtget tatgecata agaagacae 18000 gactetgace etetgatg gagacteeca agaactetta ggaacaagac acaattaagg 15000 gactettgace tecacetet agaaggeggtgggagaa attttttggg gaacagaatg 15100 tecatgtget tecaataag geatgaagag gtgtggaagaa attttttggg gaacagaatg 15100 gagggaatec acctetetea ggaagecete atetggaaaa gettaagaag gtgtgaagaga 15240 aactgtagg ecagteettg gtgtecaaaa tggggtttta ggtttggg gaacagate 15300 tecatgtage caagtettg accaatgt tegteacac acctgtaa teggaggace 15300 tecatgtage ageteettg gtgeceatgte tecacetgt teatgetee teaggageag 15240 aactgtaga ageteetta gacaagtgg etecaceagt tecaceagaa atatgacea 15420 aggattgaagg teatgagag tecacaagtg gaacaacte acctgeaag atatgacea 15400 aggattgaagg teatgagga tecacaagtga gaacaacte acaagttaa attagtacea 15600 aggttgaaggt teatgagga tecagaagtg tggggaagaa accaaggaga aacaaggaagag 15780 ataagaagaa teatgagag tecacaagtga gaacaacte agaagtgaca agaagacaa 15400 aggatgagagga caagagga teggaagagga teggagagaga acaaagagaa acaagagaagaagaaga							
aataagagtt gytggcagtg ccaggcggg tytaattoca agcacttggg 14280 aggcagaggag aggcagattt ctcagttega gyccagcottg tytaacaaaa taaaaaaaaa 14340 gacagccagg gyctatacaag gaaaccctgt gycaaaaaa caaaaaaaaa 14460 ccagaaggaa cytagggtat ataccaggtg gaatttcaa atgagggt gaagctgtaa 14580 ccagaaggaa tytggcactaca gccactgtt gaattgtgga gaacacaact cagcagcat 14580 ctacatgcata agccatgct ctaactgggaa gacacaggaa acacagatag 14680 cgggttagag tggccatggc ctttcetgec tgcctttct caagggaa acacagaaga 14680 cgggttagac taacttetc attgctetct taagggtcaa cagattagt 14760 cettgggata ctcatactt attgctetta tattgattat cagaggtcta ctcatagggttt 1480 gaggttaact acacttata attgttatta tattgttaaca aggaagaca 14940 ggattatettg tacttaatag gtattaattg <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>							
aggcagaggg aggcggattt ctgagttcga ggccagcagg tctatacaga gaaaccctgt gtcgaaaaa caaaaaaaaa aaaaaaaaaa 14400 gttggcagag cttaagaggat ataccaggtg gaagtttcaa atagaggat gaagttgggtta 14460 ccagaaggaa cttaagaggat ataccaggtg gaagttaaa 14520 attgagagag tggccatca gccactgtgt gaatgtagga gaacaaaac 14520 attgagagag tggccatgac gccactgtgt gaacacaaga acacaggaa 14580 cggggttagg tggccatggc cttctctactc taacaggad acacaggaa acacagaaga 1460 cggggttagg tggccatggc ctttctcatct taacacaggaa acacaggaga 1460 ccttgggtac atcatgtgat ctgtgtgata ccagggttaa caagagttca ttttgtac tttgttaat ccaggactcta ttgatgacaaa aggaagaca 1480 1480 ggattatettg acatctactc aactttaat acatttaat acattgagac ttatgacaaa 1500 1494 ttatttatta aattaatagga </td <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>							
gacagecagg getatacagg gacacectgt gtegaaaaa caaaaaaaa aaaaaaaaa l4400 gttggcagag tgtgggttat ataccaggtg gagattteaa atgagtggc gaagectgtag 14520 attgagagag tgtgggcacaca gecactgtgt gaaatgtggca gaacacaac cagcagaaa 14520 attgagagag tgggcacacac gecactgtgt gaattgtgca gaacacagga tagcagaaac 14520 cacaggattag tgggcacacac gecactgtgt tgactette tagcacaagga aacagataga 14640 cggggttatag tggcgcatec tgcctettec taagggtaa aacagataga 14640 cettgggtac ctaactettec atcattgtag tecttaat agttectag cttettaat actteaggtcc ctttgat gaggttacac acaaggtgaag tgattcatg aggagacaaca aggagaagca aagagttaca aggagaagca teatgcaaga atcatgggtc tatagcaaga gaaacatgg tteatgcaat acaatggtg 15900 ttecttgtte tecaattag getataatagag gtgttgtaaaa atcttgtttec agaa							
gttgggagag tgtgggttat ataccaggtg gagatttcaa atgagagga 14460 ccagaaggaa cttagaggat agctcataac ttaaaaagaa atgagagag tagcagaaac 14520 attgagaaga tgggcacaca gccactcgtgt gaatgtggca gaacacaatc cagcagaca 14580 tacatgataa agtgtatatt ggcgcatcc tgcctcttcc taacacagaa aacagataag 14640 cggggttagg tggcatcc tgcctcttcc taagggtcac ctcaagact 14700 ccttgggtac atcagtgac cttcctctcc atcgttctaac tatacctccc agattagcac 14760 ccttgggtac atcagtgac ctcatctcc atcgttcaca tgtgtgacaa attaccttcac ctgattcaa aggagagac 14820 gaggctgcaa catcatcatct actcatcatc actcatcatc actcatcatc agacactcta ggaacagac atcagggtc 14940 gaattatca actagcatca actcatcatca agacatcaca agcacatcac acacattacac 15000 ttgttttac ttccattaca gcattcata ggat							
ccagaaggaa cttagaggat agctcataac ttaaaaggaa tagagaagg tagagaagg tagagaaggaagg tagagaagaagg tagagaagaagg tagagaaagaagg tagagaaagaagg tagagaaagaagg tagagaagaagaagg tagagaagaagaaggaaggaaggaaggaaggaaggaagg							
tatagagaga tgggcacaca gccactgtgt gaatgtggca gaacacaatc cagccagcta 14580 tacatgcata agtgtatatt ggcgcatcc tgactgatga gacacaaggaa aacagataga 14640 cggggttagg tggccatggc ctttcctgcc tgcctcttcc taagggtcat ctcaagacct 14700 tatgctctct taactcttcc attgctactt agcttctaga tatcacactca cagtatagac 14700 cctgggtac atcagtgatc ctggtgatat ccagggctc ctgattccat ctttgtcata 14820 cggggttgaa ctaaagaggt cttcttaata cttcacaccc ctgattccat aggaagacaca 14880 agaagttcac agaggtgaga tgattcatgt aggacataca gtgaggcaaac acaagggtca 14940 ggattatctg actctactct aactttatg taaatgtgct ttatgccatt acaagtggtc 15000 ttcctgtgct tcagctctgg gagactccca agcactctta ggcacaagcc acaattaagg 15060 gactctgaca ctctgcattg attaattagc atggtggtc ctatgttcc agattcatg 15120 ttgtttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaactc acctctctca ggaagcctct atctgaaaa attttttggg gacagaactc acctctctca ggaagcctct atctgaaaa acttttttggg gacagaactc 15240 acctgtgggc cagctcttg gtgtccaaaa tgggttttat ggtttgaatc tggaagagct 15300 tccatgtgc caaaggtttg aacatggagc ctcctctgg taacactgta ttggaagctg ttggaagact ttggaagactg atgctcttg gtccatgtt ttgctacatc actgtcaag atgctcttt gtccatgt ttgctacatc actggaagct 15300 ttgagactga cagctcttt gtccatgat ttgctacac actggagct ttggcaaaactc actggaagact ttggagactg ttgagagaactga atgctcttt gtccatgat ttgctacac actggagact tcgaagaggact tcctcctgg taacactgta ttggaggct 15360 ttgagactga atgctcttt gtccatgat ttgctacact actgtcaagaggact tcgaagaggacaagagacaagaacact actggagact tcattgctc agcaagttga gacaaacttt tcctgcaata attagtacca 15480 tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15400 gagttgaggt tcattgctc agcaagttgg atcaaacttt tcggggctaca aggaactgaagaga actgaagaga acctaggga aataggaaa atcaaactt aggaagaga acctaggga aataggaaagaa accaaggaaagagaa accaaggaaagagaaagaa							
teacatgcata agtgtatatt ggcgccatec tgactgatga gacacaggaa aacagataga 14640 cggggttagg tggccatggc ctttcctgcc tgcctcttcc taagggtcat ctcaagacet 14700 tatgctctct taactcttcc attgctactt agcttctaga tatcacctcc agattagtc 14760 ccttgggtaa atcagtgate ctggtgatat ccagggcttc ctgattccat ctttgtcata 14820 gaggctgcaa ctaaagaggt cttcttaata cttcacaccc tgattccat agagagaca 14880 agaagttcac agaggtgaag tgattcatgt aggacaaca gtgagcaaaa aggaagacac 14880 ggattatctg actcatctt aacttttatg taaatgtgct ttatgccatt aacactgtca 15000 ttcctgtgct tcagctctg gagactccca agcactcta ggcacaagc acaattaagg 15060 gactctgaca ctctgcattg attaattagc atggtggtct ctatgttca agattcatcg 15120 ttgttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaactc acctctctca ggaagccctt atctggaaaa gcttacacct cagggacatcc acctctctca ggaagcctt atctggaaaa gcttacacat cagggacatc cagtctggtg acaacggactet gtgccaaaaa tgggtttat ggttgaatc tggaaggct 15300 tccatgtgct caaaggtttg aacatggagc ctcctctgg taacactgta ttggaaggct ttggaaggct ttggaagactg atggaagctg acaaggcttt ggccatgt ttggtaaca attggaaggct ttggaaggct ttggaaggct ttggaaggctgaactgaa							
cggggttagg tggccatggc ctttcctgc tgcctctcc taagggtca ctcaagacct 14700 tatgctctct taactctcc attgctact agcttctaga tatcacctcc agattagtc 14760 ccttgggtaa atcagtgat ctggtgatat ccagggcttc ctgattccat ctttgtcata 14820 gaggctgcaa ctaaagaggt cttcttaata cttcacaccc tgatgccaaa aggaagacac 14880 gaatatactg actcacct aacttttatg taaataggc ttatgccaaa aggaagacac 15000 ttcctgtgct tcagctctgg gagactccaa agcactctta ggcacaagc acaattaagg 15060 gactctgaca ctctgcattg attaattagc atggtggata atttttggg gacagaattg 15120 ttgtttcact tcccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatcc acctctca ggaagccct atctggaaa gcttacaact cagggacagt 15240 aactgtaggc ccaagttg aacatggagc ctcctctggg taacactggagc caagggacattg acctatggag ccaaggattg aacatggagc ctcctctgg taacactgta ttggaggata ttggagaatgg atggagaatcc acgggacagt accaggacagt ttggagaatgg atggagaat ttggagactgg atggagaact tccatggaaa gcttacaact cagggacagt 15300 tccatggtgc caaggtttg aacatggagc ctcctcctgg taacactgta ttggaggctt 15360 ttgagaatgg atgctctttg gtcccatgtt ttgctacact atctgcaag atatggagct 15480 tagacatga tctcctaaaa atggaatcaa agcaaacttt tcctgcaaga atatgacca 15420 gagactgcta cagctacaa atgggaatcaa agcaaacttt tcctgcaata attgttctc agcaagtgg acaagacatc acaacacagaa agttttttt 15540 tttctgttaa gtgtttggtc acagggacaa gaaaacactc actacagata attagtacca 15600 gagttgaggt tcattgctc agcaagttgg atcaaatttt taggggcttg gaactgatt 15600 ataagaagac tgtagaagag tctgaaggg tggggcaaca gaaaacactc acaacaaa ggcagacagg 15780 aaaacgggag caatgtggg tcattgagagg tctgaaggg tggggcaaca aacacacaaa aggacaggg 15780 aaaacgggag caatgtggg tggagaggg tggggggggc ataagaagg acctagggg aaatgggag aaatgggaa aataggca 15900 ctgaatgagg ccaaatttta aaggagtgg tggagaggc tggagagaag acctagggg aaatggaga 15900 ctgaatgagg ccaaatttta aaggaaggga tggagaagga tggaggaagagga tggagaagaga accaagagaa aatacaaga 15900 ctgaatgagg ccaaatttta aaggagatgg tggaggaa atcatggaa atcaagaa atcaaggaa atcaaggaa atcaaggaa 15900 ctgaatgagac ccaaggcaa tacaaggaa tggggagaaa atcaaggaa aatacaaga aggagacacaca ccaaggcaa tgggggaaa atcaaggaa atcaaggaa aacaacaca tgaagaacacaca ccaaggcaa a							
tatgetetet taaetettee attgetaett agettetaga tateaeetee agattagtet 14760 cettgggtae ateagtgate etigtgatat ceagggette etigteeaa detigteeta 14820 gaggetgeaa etaaagagg etietetaata etiecacaee tgatgeeaaa aggaagaeae 14880 agaagttee agaggtgaag tgatteatgt aggacataea gtgageaage ateagggtee 14940 ggattatetg actetaetet aactettatg taaatgtget titatgeeatt aacaetgtea 15000 titeetgiget teageteig gagacteeea ageaetetta ggeacaagee acaattaaga 15060 gaetetgaee etiecataga getatagaag gtgatgaggaa attititigg gacaagateg 15180 gaggeaatee acetetetea ggaageetet atetggaaga attititigg gacagaagatg 15180 gaggeaatee acetetetea ggaageetet atetggaaaa getiaeaaet eagggacagt 15240 aactgtagge ecagteetig gtgiceaaaaa tgggittiat ggittigaate tigeaagaget 15300 tigaggaatee eageteetig gidecaaaaa tgggittiat ggittigaate tigeaagaet 15300 tigagaatgga atgeteetig gidecaaaaa tgggittiat ggittigaate tigeaagaet 15300 tigagaactga atgeteetig gidecaaaaa tigggittiat ggittigaate tigeaagaet 15300 tigagaactga atgeteetig gidecaaaaa tigggittiat ggittigaate tigeaagaet 15480 tigagaactga atgeteetig gidecaatga ageaaacete teatgitee eecacacaga 15480 tigagaactga atgeteetig gidecaatga agaaacaete teatgitee eecacacaga 15600 gagttgaggt teatiggte acaaggaacaa gaaaacaete aateagaata attagacaa 15600 gagttgaggt teatiggete agaaggtgg ateaaactee teatgitee gaactgatet 15500 aataagagaa tigagaagag tetgaaagag ateaaacte taagggettig gaactgatet 15600 aataagagaa tigagaagag tetgaaagaga ateaaaactee agaatgataa ateagaaga 15780 aataagatgaa eatigggget gigagaagaga ateagaagaa acetaggagg aaatgagaaga 15780 aataagagga cateagget gigagaagaga ateacagaa gigagaagaga acetagggga aataacaagaa 15940 etigaatgaa eatigaggaa taacagaga tigagaagaga etaacacaa gigaaagaa 15840 etaagaacacaca etaaggeta geggaacggi gigageggaa aateagaga aataacaaga 15960 etaagaagaga cataaggaa taacagaaa gegaacacaca etaaggeta geggaacggi gigagagaa acetagggga aataacaagaa 15940 etagaacacaca etaaggeta geggaacgga tigagaagaa acetaggaga aataacaagaa 15960 etaagaacacaca etaaggeta taacaagaa taacacagaa taacacacaca taacagacaa taacagacaa ateacacaca agaacacacaca ateagagaa t							
cettgggtaa atcagtgate etggtgatat ecagggette etgatecat etteteata 14820 gaggetgeaa ctaaagaggt ettettaata ettecacaeee tgatgecaaa aggaagacae 14880 agaagteca agaggtgaag tgatteatgt aggacataca gtgageaage atcagggtee 14940 ggattatetg actetactet aactttatg taaatgtget tatgecatt acacatgtea 15000 tteetgget teagetetg gagactecca agcactetta ggeaaagee acaattaagg 15060 gactetgaca etetgeattg attaattage atggtgget etatgttee agatteatgg 15100 ttgtteact tecacataag getatgaagg gtggaggaa atttttegg gacagaattg 15180 gaggeaatee acctetetea ggeageetet atetggaaaa gettacaaae eagaggacagt 15200 tecatgtget ecagteetig gtgtecaaaa tggggttta ggtttgagga atttettggg gacagaattg 15180 acctgtagge ecagteetig gtgtecaaaa tggggttta ggtttgaget eagaggacaet eagaggacaet eacatggage ecagteetig gtgtecaaaa tgggtttat ggtttgaate tggaaggett 15300 tecatgtget ecaaaggettg acaatggage etceteetig taacactgta ttggaggett 15300 tecatgtget eaaaggtttg acaatggage etceteetig taacactgta ttggaggett 15300 tetgagactga atgetettig gteceatgt ttggetacate acetgtaaga atatgacee 15420 ggeatgetae eagetacae agactatgee teceagett tecatgetea attggagett 15540 ttggagaetga teceeaaa atggaateaa agaaacaete tecegeatt aggattgaggt tecatgete ageaagttgg atcaaatttt tagggetttg gaactgatt 15600 gagttgaggt tecatgete ageaagttgg atcaaatttt tagggetttg gaactgatt 15600 ataaggagae tggaagagg tetgaagetg tggggetacag aagtgeeae agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg etcacacaaa ggeagacagg 15780 aaaacgtgag catgtggget gtggaaggge ataagaagga ecaaactta agggaaggge etaactgaaggg eagacacaca etcaggeta gggaagggg etaactgat tggeeaagggg eagacacaca etcaggeta gegaagggga etaactgaa tacacgata tgeagacacaca etcaggeta gegagaggga etaactgaa atecagata atecagaa aatggaacacaca etcaggeta gegagaggga etaactgaa atecaggagg eagacacaca etcaggeta gegagaggga etaacacaca tgaggeaa atecaggaga aatggagaa atecaggaga aatggaacacacacacacacacacacacacacacacacac							
gaggetgaa etaaagaggi ettettaata etteaacee tgatgecaaa aggaagacae 14880 agaagtteae agaggtgaag tgatecatgi aggacataca gtgagcaage accagggtee 14940 ggattatetg actetactet aactittatg taaatgtget tatgecatt aacactgtea 15000 tteetgget teagetetg gagaceeca ageactetta ggeacaagee acaattaagg 15060 gactetgaca etetgatg attaattage atggtgget etatgttee agatteatga 15120 ttgttteaet teecatatag getatgaagg gtgtgaggaa attitttggg gacagaattg 15180 gaggeaatee acctetetea ggaageetet atetggaaaa gettacaact eagggacagt 15240 aactgtagge ecagteettg gtgteeaaaa tgggttttat ggtttgaate tgegaaggeet 15300 teeatgtget eaaaggttig aacatggage eteeteetig taacactgta ttggaggeet taggaggaatgeet eaggacaget 15300 teeatgtget eaaggetetig geeeaaggeet eteeteetig taacactgta ttggaggeet 15300 tegagaatgga atgetetitg gieceaatge eteeteetig taacactgta ttggaggeet 15300 tegagaatgga atgeteetig gieceaatgee eteeteetig taacactgta ttggaggeet 15300 tiggagaatgga atgeteetig gieceaatgee eteeteetig taacactgta ttggaggeet 15300 tiggagaetig atgeteetig gieceaatgee eteeteetig taacactgaa atatggacea 15400 gagattgaa teeteetigaa aagaataaa agaaaacte eeteetigaa atatgacea 15400 titletgaa gtgtttggte acaggagaaa gaaaacacte aatacagata attagtacea 15400 gagttgaggt teattgeete ageaagtgga acaaaacacte aatacagata attagtacea 15600 gagttgaggt teattgeete ageaagtgga ataaaattit tagggettig gaactgatti 15600 aaaaacgtgag eatggegga gaaatgggaa aatagggaa actaggggg aaatggaca aggaacgatt 15700 aaaaacgtgag eatgggegg gtgggagggeaaggaa aatagggaa aatagggaa aataggaa aataggaa aataggaga eetgaaggga eatagggga aatagggaa aatagggaa aataggaca 15700 etgaatggag eatagggea gggaacggtg gtgggetgg ttggeecata ggeagaacggi 15780 aaaaacgtgag eatagggea gggaacggtg gtgggetgg ttggeecata ggeagaacgg 15780 aaaacgtgag eatagggaa eatagggaa aatagggaa aataggaa aatacaaga 15900 etgaatgagg eaaaattta aaggaggtga tgaagacatt tgeaggaaa aatacaaga 15900 etgaatgagge eetaaggea tgggaacgat tateagagga tggagaaca tacaggaa aatacaaga 15900 etgaatgagge eetaagagga tggggaaca aatacaaga 15900 etgaatgagge eetaaggaga tggggaacgate tgaagaata aatacaaga aacacactt 16000 ggtagggte eetaaggagga tgg							
agaagttcac agaggtgaag tgattcatgt aggacataca gtgagcaagc atcagggtcc 14940 ggattatctg actctactct aacttttatg taaatgtgct ttatgccatt aacactgtca 15000 ttectgtgct teagetetgg gagactecca agcactctta ggcacaagcc acaattaagg 15060 gactctgaca etctgcattg attaattage atggtggtc ctatgttcc agattcatga 15120 ttgtttcact tteccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatcc acctctca ggaagcctct atctggaaaa gcttacaact cagggacagt 15240 aactgtaggc ccagtcettg gtgtccaaaa tgggttttat ggtttgaatc tgcaaagcct 15300 tccatgtgct caaaggtttg aacatggagc ctcctcctgg taacactgta ttggaggct 15360 ttgagactgg atgctctttg gtcccatgtt ttgctacact atctgcaag atatgaccca 15420 ggcatgctac cagctaccac agactatgcc tctccagctt tcatgttctc cccacatga 15480 tagacttgta tctcctaaaa atgggactac ctcccagctt tcctgcata agttttttt 15540 tttctgttaa gtgtttggt acagggacaa gacaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctc aggaggacaa gacaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctc agcaggtaga atcaaattt tagggcttg gaactgattt 15660 ataagagaca tgtagaagag tctgaaggtg gtgggctacag aagtgtcacc aggttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780 aaaacgtgag catgtgggct gtgagagggc ataagaagga acctaggggg aaattgagaa 15900 ctgaatgagg ccaaatttta aaggagtgg gtgaggagaa atcaactca agcaacacca ctcaggctat gcggaacgtg gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg cccaaatttta aaggagtgga ctaacctcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccaaggcaa tatcaaggaa tgtgtgcataa atcatgtag gtgatgccta agtgaccatt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacatt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa aggaaggcat tgcagtcagg 16320 agccaaaaggt cagttgcag gcaacacttc agtgacacact aggaaggcta aggaccacacacacacacacacacacacacacacacac							
tccttgtget tcagctetgg gagactecca agcactetta ggcacaagce acaattaagg 15000 gactetgaca ctettgcattg attaattage atggtggtet ctagttecca ggcacaagce acaattaagg 15100 ttgtttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatce acctetcca ggaagcetct atctggaaaa gcttacaact cagggacagt 15240 aactgtagge ccagtcettg gtgtccaaaa tgggttttat ggttgaatc tggaagget tcatagtgg accagttg accatggagg ctccttgg taacactgga atggttgaatc tggaagactg accatgtgg accagtttg gtcccatgtt ttgctacact atctgcaag atatgaccca 15420 ggcatgctac cagctacca agactatge tcccatggt tcatagattg accatgga atggaatca agcaaacttt tcctgaaga attagaccca 15420 ggcatgctac cagctacca agactatge tcccaagett tcatggatta aggattgagg tcattggat acaagggacaa accatgga atatgacca agcaaacttt tcctgaata agtttttt 15540 tttctgttaa gtgtttggtc acaagggacaa gaaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctc agcaagttgg atcaaagttg accaagttgg atcaaagttg acaaggagga accaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaaggagga acaagacacca cagaggaaa tacaggaaa ggaaggaga acaaggagga acaagacacaca aggaagga							
treetgtget teagetetgg gagaetecea ageaetetta ggeacaagee acaattaagg 15060 gactetgaca etetgeattg attaatage atggtgget etatgttee agatecatga 15120 ttgttecaet teecatatag getatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gagggeaatee aceteteca ggaageetee atetggaaaa gettacaact cagggacagt 15240 acetgtgagge ecagetettg gtgtecaaaa tgggttttat ggtttgaate tggcaaageet 15300 teeatgtget caaaggtttg acetggage eteetetgg taacactgta ttggagget 15360 ttgaggactgg atgetettg gteceatgt ttgetacate atetgteagg ataggeet 15420 ggcatgetae eagetaceae agaetatgee teteceaget teatgttee eccacatga 15480 tagacttgta tetectaaaa atggaateaa ageaaactee teetggatta agtttttt 15540 tttetgttaa gtgtttggte acagggacaa gacaacete teetggatta agtttttt 15660 gagttgaggt teattgete ageaagttgg ateaaatte tagggetttg gaactgatt 15600 gagttgaggt teattgete ageaagttgg ateaaatte tagggetttg gaactgatt 15600 gagttgagg teatggagg tetgaaggg ateagaagtg tegggetacag aateaggata attagtacea 15720 aataagttaa tacaccatgg gaattgtgaa aateagaatg eteaaagatg etgagaggg aataagagg etgagaggg aataggagg etgagaggg etgagaggg etaaacatee ggeaacacaca etcaggetg gtgagaggge gtgagaggg etaacegaa acetaggggg aatagaagg 15780 etgaageegg etaagggg etaagaggg etaagagggg etaagagggg etaagagggg etaagagggg etaagaggg etaagaggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggg etaagaggggg etaagagggg etaagagggg etaagagggg etaagagggg etaagagggggggggg							
ttgtttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatcc acctctctca ggaagcctct atctggaaaa gcttacaact cagggacagt 15240 aactgtaggc ccagtccttg gtgtccaaaa tgggtttat ggtttgaatc tgcaaagcct 15300 tccatgtgct caaaggtttg aacatggagc ctcctcctgg taacactgta ttggaaggct 15300 ttgagactgg atgctctttg gtcccatgtt ttgctacatc atctgtcaag atatgaccca 15420 ggcatgctac cagctaccac agactatgcc tctccaggt tcatggtct ccacacatga 15480 ttgagactgta tctcctaaaa atggaatcaa agcaaacttt tcctgcaag atatgaccca 15420 ggcatgctac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480 ttgagacttgta tctcctaaaa atggaatcaa agcaaacttt tcctgcatta agttttttt 15540 tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660 ataaggagaca tgtagaagag tctgaaggtg atcaaatttt tagggctttg gaactgattt 15660 ataaggagaca tgtagaagag tctgaaggg atcaaaattt tagggctttg gaactgattt 15600 aaaacgtgag catgtgggct ggtgagagggc ataagaaga acctaagggg aaatgagaa 15780 aaaacgtgag catgtggcgt gtgagagggc ataagaaga acctaggggg aaatgagaa 15840 gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatacaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccaagagcaa taccagaca tgtggaata atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggaa aggaaggcat tgcagaggt aggaccaaaggt tgcagggtc aggacaaggt tgcagggtc aggacggtca aggaaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgagaag gccagaaggtca aggaaggca tgcagaaggt tgcaggaggcaaggca							
ttgtttcact ttccatatag gctatgaagg gtgtgaggaa attttttggg gacagaattg 15180 gaggcaatcc acctctctca ggaagcctct atctggaaaa gcttacaact cagggacagt 15240 aactgtaggc ccagtccttg gtgtccaaaa tgggttttat ggtttgaatc tgcaaagcct 15300 tccatgtgct caaaggtttg aacatggagc ctcctctgg taacactgta ttggaaggct 15360 ttgagactgg atgctctttg gtcccatgtt ttgctacatc atctgtcaag atatgaccca 15420 ggcatgctac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480 tagacttgta tctcctaaaa atggaatcaa agcaaacttt tcctgcatta agttttttt 15540 tttctgttaa gtgtttggtc acagggacaa gaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660 ataaggagaca tgtagaagag tctgaaggtg atcaaatttt tagggctttg gaactgattt 15600 aaaacgtgag catgtgggt gggaggacag aatacagaa agtgtcacc agttttaag 15780 aaaacgtgag catgtgggt gtgagagggc ataagaagg acctaggggg aaatgggag aataggaga 15780 aaaacgtgag catgtgggt gggaacgtgt gtggctgtgc ttggcccatg ggcaacagg 15780 cagaccacca ctcaggcta gcggacgtgg ctaacctgat tgcaggaga acctagggg aaatgagca 15900 ctgaatgag ccaaatttta aaggagtgga ctaacctgat tgtcagagaa aatatcagaa 15900 ctgaatgagg ccaaatttta agggatgga ctaacctgat tgtcagagaa aatatcagaa 15900 ctgaatgagg ccaaatttta accgagcta taccagaca tgaagacaa tacaggttag gtgatgggc cctaagtgga tggtggataa atctatgtag gtgatgccta agtgacctt 16020 ttgtgaaatt ccaagagcaa taccagaa atgtggaata tcttccaagg agacctgtag accacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag accacactt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16200 gtctgagtta cggttccagg gcaacacttc agggtagaga gccagaggtta tgcagaggt aggaccaaggt tgcagaggt aggaccaaggt tgcagggta aggaccaaggt tgcagaggt aggaccaaggt tgcagaggta aggaccaaggt tgcagaggt aggaccaaggt tgcagaggt aggaccaaggt							
aactgtagge ccagtecttg gtgtccaaaa tgggttttat ggtttgaate tgcaaaget 15300 tecatgtget caaaggtttg aacatggage ctcetetgg taacactgta ttggaggett 15360 ttgagactgg atgetetttg gtcccatgtt ttgetacate atetgtcaag atatgaceca 15420 ggcatgetae cagetaceae agactatgee tetecagett teatgtete eccacetga 15480 tagaettgta tetectaaaa atggaateaa ageaaacttt teetgeatta agttttttt 15540 tttetgttaa gtgtttggte acagggacaa gaaaacacte aatacagata attagtacea 15600 gagttgaggt teattgete ageaagttgg ateaaacttt tagggetttg gaactgattt 15660 ataaggacaa tgtagaagag tetgaagetg tgggetacaag aagtgtcace agtttttaag 15720 aatagttaa tacaccatgg gaattgtgaa aatacagaatg eccacaaaa ggcagacagg 15780 aaaacgtgag catgtggegt gtgagaggge ataagaaga acctaggggg aaatggaag ecaaactte ggetacgtta gggaacgtgt gtgggetgge ttggeccatg ecetgggaat 15900 etgaatgagg ecaaattta aaggagtgga etaactegat tgaagagaa aatacaaga 15900 etgaatgagg ecaaattta aaggagtgga etaactegat tgaagagaa aatacaaga 15900 etgaatgagg ecaaattta aaggagtgga etaactegat tgaagaaca tacagtttag tgaagtaag 16080 ggttgggtt ectaaggga tggtggaatg tetacaaga ateatgtag gtgatgecta agtgaacett 16140 gataatccaa aatacagaa atgtggaatg tetacaaga ateatteet gaaggagtta 16200 tagaactttg etcatggetg taataaatag etagetagaa accattteet gaaggagtta 16200 tagaactttg etcatggetg taataaatag etagetagaa ateatteet gaaggagtta 16200 tagaactttg etcatggetg taataaatag etagetagaa ateatteet gaaggagtta 16200 gettgagtta eggttecagg geaaacatte aggtaggaa aggaaggeta tgeaggatta 16200 gettgagtta eggttecagg geaaacatte aggtaggaa aggaaggeat tgeagtcagg 16320 gettgaggtta eggttecagg geaaacatte aggtaggaa aggaaggeat tgeagtcagg 16320 agecaaaaggt eagetggte eactggtea eattgeatea aggaagaggaa tgeagaagget 16320 gettgaggtta eggttecagg geaaacatte aggaagaggaa aggaaggeat tgeaggaaggeaggeaggeaggeaaggea							
tccatgtgct caaaggtttg aacatggage ctcctctgg taacactgta ttggaggctt 15360 ttgagactgg atgctctttg gtcccatgtt ttgctacate atctgtcaag atatgaccca 15420 ggcatgctac cagctaccac agactatgce tctccagett tcatgttcte cccaccatga 15480 tagacttgta tctcctaaaa atggaatcaa agcaaacttt tcctgcatta agttttttt 15540 tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660 ataagagaca tgtagaagag tctgaaggtg tggggtacaag aagtgtcacc agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780 aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagca 15940 gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tctccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcattccc gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggaa aggaaggcat tgcagtcag 16320 agccaaaggt cagctggtca cattgcatca agtgatgaa gtcagagtt gagtagaaag 16380							
ggcatgctac cagctaccac agactatgcc tetecagett teatgteca atetgtecag atatgacca 15420 ggcatgctac cagctaccac agactatgcc tetecagett teatgtect eccaccatga 15480 tagacttgta tetectaaaa atggaatcaa agcaaacttt teetgcatta agttttttt 15540 tttetgttaa gtgtttggte acagggacaa gaaaacacte aatacagata attagtacca 15600 gagttgaggt teattgetet agcaagttgg atcaaatttt tagggetttg gaactgattt 15660 ataagagaca tgtagaagg tetgaagetg tgggetacag aagtgteace agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg etcacacaaa ggcagacagg 15780 aaaacgtgag catgtggegt gtgagaggge ataagaagga acetaggggg aaatgageta 15840 gaagecatte ggetacgtta gggaacgtgt gtggetgtge ttggeccatg ecetggcaat 15900 etgaatgagg ecaaatttta aaggagtgga etaactegat tgtcagagaa aatatcaaga 15960 eagaceacca etcaggetat geegtgtttg tgacegacca getaetetta geeageteta 16020 ttgtgaaatt ecagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggte ectaatggetg tagtggaatg tettecaagg agacetgtag acacacattt 16200 tagaactttg etcatggetg taataaatag etagetagaa atcattect gaagaggtta 16260 gtetgagtta eggttecagg geaaacatte agtgatggaa ggcagaggtt tgeagtcagg 16320 agecaaaggt eagetggtea cattgeatea aggaaggaa gteagaggtt gagtagaaag 16380	aactgtaggc	ccagtccttg	gtgtccaaaa	tgggttttat	ggtttgaatc	tgcaaagcct	15300
tagacttgta tetectaaaa atggaateaa ageaaacttt teetgeatta agttttttt 15540 tttetgttaa gtgtttggte acagggacaa gaaaacatt teetgeatta agttttttt 15540 tttetgttaa gtgtttggte acagggacaa gaaaacatt teetgeatta agttttttt 15540 gagttgaggt teattgetet ageaagttgg ateaaatttt tagggetttg gaactgattt 15660 ataagagaca tgtagaagag tetgaagetg tgggetacag aagtgteace agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aateagaatg eteacacaaa ggeagacagg 15780 aaaacgtgag catgtggegt gtgagaggge ataagaagga acetaggggg aaatgageta 15840 gaagecatte ggetacgtta gggaacgtgt gtggetgtge ttggeecatg eeetggeaat 15900 etgaatgagg ccaaatttta aaggagtga etaactegat tgteagagaa aatateaaga 15960 cagaccacca eteaggetat geegtgtttg tgaecgacca getactetta geeageteta 16020 ttgtgaaatt eeagageaat tateagagea tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggte eecaaggetg taataaatag etatecaag ageactgtag acacacattt 16200 tagaactttg eteatggetg taataaatag etagetagaa ateatteet gaagaggta 16260 gtetgagtta eggttecagg geaaacatte agtgatggaa aggaaggeat tgeagtgaag 16320 ageeaaaggt eagetggtea cattgeatea agtgatggaa ggteagagtgt gagtagaaag 16380	tccatgtgct	caaaggtttg	aacatggagc	ctcctcctgg	taacactgta	ttggaggctt	15360
tttetgttaa gtgtttggte acagggacaa gaaacacte aatacagata attagtacca 15600 gagttgaggt teattgetet agcaagttgg atcaaatttt tagggetttg gaactgattt 15660 ataagagaca tgtagaagag tetgaagetg tgggetacag aagtgteace agttttaag 15720 aatagttaa tacaccatgg gaattgtgaa aatcagaatg etcacacaaa ggcagacagg 15780 aaaacgtgag catgtggegt gtgagaggge ataagaagga acctaggggg aaatgagca 15840 gaagecatte ggetacgtta gggaacgtgt gtggetgtge ttggeecatg ecetggeaat 15900 etgaatgagg ecaaatttta aaggagtgga etaactegat tgtcagagaa aatatcaaga 15960 eagaceacca etcaggetat geegtgtttg tgacegaca getactetta geeageteta 16020 ttgtgaaatt ecaagtgga tggtgcataa atetatgtag gtgatgeeta agtgacactt 16140 gataatecaa aatatcaga atgtggaatg tettecaagg agacetgtag acacacatt 16200 tagaactttg etgatecag geaacatte agtgataga ateatteet gaagaggtta 16260 gtetgagtta eggttecagg geaacatte agtgatgga aggaaggeta tgeagaggt aggeeaaaggt eagetggta aggaagagga gteagaggt gagtagaaag 16380	ttgagactgg	atgctctttg	gtcccatgtt	ttgctacato	atctgtcaag	atatgaccca	15420
tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600 gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660 ataagagaca tgtagaagag tctgaagctg tgggctacag aagtgtcacc agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780 aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840 gaagccattc ggctacgtta ggggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggaa ggcagagtgt gagtagaaag 16380	ggcatgctac	cagctaccac	agactatgcc	tctccagctt	tcatgttctc	cccaccatga	15480
gagttgaggt teattgetet ageaagttgg ateaaatttt tagggetttg gaactgattt 15660 ataagagaca tgtagaagag tetgaagetg tgggetacag aagtgteace agttttaag 15720 aatagttaa tacaccatgg gaattgtgaa aateagaatg etcacacaaa ggeagacagg 15780 aaaacgtgag catgtggegt gtgagaggge ataagaagga acetaggggg aaatgageta 15840 gaagceatte ggetacgtta gggaacgtgt gtggctgtge ttggeecatg eeetggeaat 15900 etgaatgagg ecaaatttta aaggagtgga etaactegat tgteagagaa aatateaaga 15960 eagaceacea etcaggetat geegtgtttg tgacegacea getaetetta geeageteta 16020 ttgtgaaatt ecagageaat tateagagea tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggte ectaagtgga tggtgeataa atetatgtag gtgatgeeta agtgacactt 16140 gataateeaa aatateagea atgtggaatg tetteeaagg agacctgtag acacacattt 16200 tagaactttg etcatggetg taataaatag etagetagaa ateatteet gaagaggtta 16260 gtetgagtta eggtteeag geaaacatte agtgatggaa ggeagaggeat tgeagteaga 16380 ageeaaaggt eagetggtea eattgeatea agagtagaag gteagagtg gagtagaaag 16380	tagacttgta	tctcctaaaa	atggaatcaa	agcaaacttt	: tcctgcatta	agttttttt	15540
ataagagaca tgtagaagag tctgaagctg tgggctacag aagtgtcacc agttttaag 15720 aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780 aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840 gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380	tttctgttaa	gtgtttggtc	acagggacaa	gaaaacacto	: aatacagata	attagtacca	15600
aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780 aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840 gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtg gagtagaaag 16380	gagttgaggt	tcattgctct	agcaagttgg	atcaaatttt	: tagggctttg	gaactgattt	15660
aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840 gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtg gagtagaaag 16380	ataagagaca	tgtagaagag	tctgaagctg	tgggctacag	g aagtgtcacc	: agtttttaag	15720
gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900 ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtt gagtagaaag 16380	aatagtttaa	tacaccatgg	gaattgtgaa	aatcagaatg	g ctcacacaaa	ggcagacagg	15780
ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960 cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtgt gagtagaaag 16380	aaaacgtgag	, catgtggcgt	gtgagagggc	ataagaagga	acctaggggg	g aaatgagcta	15840
cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020 ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtgt gagtagaaag 16380	gaagccatto	ggctacgtta	gggaacgtgt	gtggctgtgc	ttggcccatg	g ccctggcaat	15900
ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080 ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtgt gagtagaaag 16380	ctgaatgagg	ccaaatttta	aaggagtgga	ctaactcgat	tgtcagagaa	a aatatcaaga	15960
ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140 gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200 tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaag gtcagagtgt gagtagaaag 16380							
gataatccaa aatatcagca atgtggaatg tettecaagg agacetgtag acacacattt 16200 tagaactttg etcatggetg taataaatag etagetagaa atcattteet gaagaggtta 16260 gtetgagtta eggttecagg geaaacatte agtgatggea aggaaggeat tgeagteagg 16320 ageeaaaggt eagetggtea eattgeatea agagtagaag gteagagtgt gagtagaaag 16380							
tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260 gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380							
gtctgagtta eggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320 agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380	gataatccaa	a aatatcagca	atgtggaatg	tcttccaag	g agacctgtag	g acacacatt	16200
agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380							
agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380							
	agccaaagg	t cagctggtca	cattgcatca	a agagtagaga	a gtcagagtg	t gagtagaaa	g 16380
aggatacagg ttataaaacc tcactgtcca ctctcagcaa tccattttct cctaaaaggc 16440	aggatacag	g ttataaaacc	tcactgtcca	a ctctcagca	a tccattttc	t cctaaaagg	c 16440

tttaccttct	aaagatttta	gtcttcaaaa	ccagtaccag	tagcctggga	acaaaagttg	16500
aaacaaatga	gcctttgtgg	ggcatttcac	acttaaaaca	gggcatcacc	taggaggagc	16560
cctgtgtgca	gtaggaagtg	tggcctctgt	gtcaggaatg	ctcaggctaa	taaggggtcc	16620
tctatctgag	ggaccctatg	aagattcaac	aagtagttgt	gagaattccc	tgtaaatgga	16680
tgctaccaat	ttgacatttg	tagacctgct	attgtgtgct	tctttattgg	gctctcccat	16740
ctcccaactt	tccaacccat	attccacatt	aatcccttcc	accaccatgc	aacactaggt	16800
aggagagaag	gaaggttaga	agagaaagtg	ggtatagatc	tatttagact	acttcctgct	16860
gattaggggc	aagtccaatc	gtcattgtca	ggatacctcc	aaccagcaac	cagcaaacca	16920
gcaaatcaga	aacagcaaaa	gcagccaaca	aggcagcact	aaccagcagg	attggggtcg	16980
gtagcgtggg	agcagtcact	actggtcttc	tcatggcttt	ggcattaata	ctctctcaag	17040
aaattccgta	${\tt atttttccc}$	caccacctga	aattccgtaa	ttttaaatgc	aaactatcta	17100
cagctggcaa	aaatcacatc	tctcctagag	cacaagacaa	atcatagtta	ctggctattt	17160
gcaatctgaa	gcatctcaat	atcccacacc	tgggattaaa	acaaaaacat	attcacatca	17220
cataactgtt	ttttttcc	aatttttat	taggtatttt	ctttatttac	atttcaaatg	17280
ctatcccgaa	agtcccctat	acceteceae	ctccctgctc	ccctacacac	ccactcccac	17340
tttttgaccc	tggagttccc	cggtactggg	gcatataaag	tttgcaagac	caaggggcct	17400
ctcttcccag	tgatggccga	ctaagccatc	ttctgctaca	tatgcagata	gagacacgag	17460
ctctgggggt	actagttagt	tcatattgtt	gttccaccta	tagggtcgca	gaccccttca	17520
gctccttggg	tactttgtct	agctcctcca	ctgggggctc	tgtgttttat	ctaatagatg	17580
actgtgagca	tccacttctg	tatttgacag	gcactggcct	agcgtcacat	gagccagcta	17640
tatcagggtc	ctttcagcaa	aaccttgctg	gcatgtgcaa	tagtgtctgc	gtttggtggt	17700
tgattatggg	atggatccac	tagttctaga	gc			17732

Tonal Application No.

PCT/CA 00/00430 A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A01K67/027 C12N9/16 IPC 7 According to International Patent Classification (IPC) or to both national dessification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 A01K C12N IPC 7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) STRAND, EPO-Internal, WPI Data, PAJ, MEDLINE, BIOSIS, EMBASE C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Category * Relevant to claim No. Y WO 97 48812 A (UNITED KINGDOM GOVERNMENT 1-69 ;YANKE LINDSEY JAY (CA); CHENG KUO JOAN) 24 December 1997 (1997-12-24) page 18, line 27 - line 31 WO 99 17610 A (UNIV CALIFORNIA) X 1-5, 10, 15 April 1999 (1999-04-15) 50-52,62 page 7, line 12 - line 26 6-9 page 10, line 21 -page 11, line 22 Further documents are listed in the continuation of box C. X Patent family members are listed in annex. Special categories of cited documents: "T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the International "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone filing date document which may throw doubts on priority ctalm(s) or which is cited to establish the publication date of another "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such document is combined with one or more other such documents, such combination being obvious to a person sidiled in the art. citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 29 September 2000 13/10/2000 Name and mailing address of the ISA Authorized officer European Patent Office, P.B. 5618 Patentioan 2 NL - 2280 HV Rijswijk Tel (431-70) 340-2040, Tx. 31 651 epo ni, Fac: (431-70) 340-3016

1

Sprinks, M

Inte :onal Application No PCT/CA 00/00430

21,22, 26,28, 29,50 15,30, 39,40, 45,46, 48,49 50,62 11–20, 30,31, 36,37, 39–49, 53–61, 63–69
26,28, 29,50 15,30, 39,40, 45,46, 48,49 50,62
39,40, 45,46, 48,49 50,62 11-20, 30,31, 36,37, 39-49, 53-61,
11-20, 30,31, 36,37, 39-49, 53-61,
11-20, 30,31, 36,37, 39-49, 53-61,
30,31, 36,37, 39-49, 53-61,
1 03 03
56,63,65
1-7, 10-14, 16-69

1

tritr ional Application No PCT/CA 00/00430

(Continue	tion) DOCUMENTS CONSIDERED TO BE RELEVANT	
egory .	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	DATABASE EMBL 'Online! ID: BLCAT3DNA, 26 February 1992 (1992-02-26) LUCKOW: "Plasmid pBLCAT3 gene for beta-lactamase and CAT gene for chloramphenicol acetyltransferase" XP002148837 Relevant for SEQ ID NO:2,6 abstract	47,56, 63,65
	DATABASE EMBL 'Online! ID: MMU73190, 4 November 1996 (1996-11-04) LAURSEN ET AL.: "Mus musculus positive acting regulatory region of the parotid secretory protein (PSP) gene" XP002148838 Relevant for SEQ ID NO:7 abstract	1-7, 10-14, 16-69
	DATABASE EMBL 'Online! ID: CVCATBLA, 24 January 1992 (1992-01-24) BOSHART ET AL.: "Cloning vector pBLCAT6 encoding chloramphenicol acetyltransferase (CAT) and beta-lactamase (bla) genes" XP002148839 Relevant for SEQ ID NO:3,5 abstract	14,20, 31,36, 44,47, 57-61, 64-69
ſ	DATABASE EMBL 'Online! ID: RNRP15, 12 May 1991 (1991-05-12) LIN ET AL.: "Rat salivary proline-rich protein (RP15) gene" XP002148840 Relevant for SEQ ID NO:3,5 abstract	1-5, 8-14, 16-69
Y	DATABASE EMBL 'Online! ID: ECAPPAA, 5 February 1991 (1991-02-05) DASSA ET AL.: "E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene" XP002148841 Encodes a phytase - relevant for present SEQ ID NO:3,5,7 abstract	11-20, 30,31, 36,37, 39-49, 53-61, 63-69

1

information on patent family members

tnti Jonal Application No PCT/CA 00/00430

Patent document cited in search report		Publication date	Patent family member(s)		Publication date	
WO 9748812	Α	24-12-1997	US	5939303 A	17-08-1999	
NO 37 1001E			US	5985605 A	16-11-1999	
			AU	716845 B	09-03-2000	
			AU	3021697 A	07-01-1998	
			BR	9709793 A	11-01-2000	
			CA	2257101 A	24-12-1997	
			EP	0904385 A	31-03-1999	
			NO	985804 A	10-02-1999	
			PL	330532 A	24-05-1999	
WO 9917610	Α	15-04-1999	US	6004944 A	21-12-1999	
MU 3317010	••		AU	9668998 A	27-04-1999	
			EP	1030557 A	30-08-2000	
WO 9908539	Α	25-02-1999	US	5876997 A	02-03-1999	
#0 1200333	••	20 02 1999	AU	8904498 A	08-03-1999	
			EP	1003379 A	31-05-2000	
			ūs	6110719 A	29-08-2000	